

STIC-Biotech/ChemLib

160642 160624

From: Ibrahim, Medina A.
Sent: Tuesday, July 26, 2005 6:18 PM
To: STIC-Biotech/ChemLib
Subject: 09/765, 555

Please search the following:

Please search the following:

1. DNA encoding SEQ ID NO: 14- 18. Thanks

Medina A. Ibrahim
Patent Examiner, GAU-1638
transgenic plants and plant breeding
Remsen-2B07
mailbox-2C18
(571) 272-0797

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacacacgaatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	778	23.6	794	5	BQ751936	BQ751936 EST632499
2	770.6	23.4	779	7	CK394523	CK394523 hggada4D0
3	768	23.3	769	7	CK118014	CK118014 218n06.p1
4	762	23.1	793	5	BQ751655	BQ751655 EST632218
5	755.4	22.9	806	5	BQ751220	BQ751220 EST631783
6	684.8	20.8	689	6	CB863541	CB863541 HH04A08Y
7	676.4	20.5	712	1	AL038548	AL038548 DKF2P566E
8	676	20.5	710	8	AY080106	AY080106 AY080106
9	647.4	19.6	660	1	AL044483	AL044483 DKF2P434I
10	641	19.4	641	4	BJ063872	BJ063872 BJ063872
11	615.8	18.7	624	6	CB863814	CB863814 HH07B20Y
12	612	18.5	794	1	AL045353	AL045353 DKF2P434B
13	603	18.3	752	9	CL655575	CL655575 PRI01248
14	589.8	17.9	608	8	AY080094	AY080094 AY080094
15	586.4	17.8	589	1	AL037742	AL037742 DKF2P564C
16	581.2	17.6	601	1	AV594538	AV594538 AV594538
17	579	17.5	579	1	AL043868	AL043868 DKF2P434C
18	570	17.3	782	7	CK394246	CK394246 hggada1B1
19	569	17.2	774	7	AL045337	AL045337 DKF2P434A
20	563.2	17.1	585	9	CL658017	CL658017 PRI0130b
21	562.6	17.0	723	1	AL039416	AL039416 DKF2P434I
22	549.4	16.6	753	1	AL045341	AL045341 DKF2P434B
23	546.8	16.6	782	1	AL037051	AL037051 DKF2P564I
24	545.4	16.5	547	6	CA895956	CA895956 B0195F10-

C 25	545	16.5	648	7	CK394349	CK394349 hggada2D0
C 26	540.2	16.4	569	1	AL039077	AL039077 DKF2P566G
C 27	534.4	16.2	644	6	CB862524	CB862524 HH04B08U
C 28	534	16.2	734	1	AL048427	AL048427 DKF2P588H
C 29	529	16.0	740	1	AL042909	AL042909 DKF2P434J
C 30	528	16.0	720	6	CA881974	CA881974 K0994C10-
C 31	523.2	15.9	528	1	AL038811	AL038811 DKF2P566O
C 32	519.4	15.7	521	6	CA890032	CA890032 B0156F03-
C 33	519	15.7	607	6	CB862099	CB862099 HH06A02Y
C 34	516.4	15.6	518	6	CA886082	CA886082 B0124F04-
C 35	510.2	15.5	686	1	AL044407	AL044407 DKF2P434D
C 36	509.4	15.4	511	6	CA895436	CA895436 B0192A07-
C 37	498	15.1	498	1	AL039076	AL039076 DKF2P566G
C 38	494.8	15.0	1025	1	AL038025	AL038025 DKF2P566C
C 39	487.6	14.8	498	7	CK394397	CK394397 hggada2H1
C 40	477.6	14.5	491	6	CB862715	CB862715 HH03F14U
C 41	465.8	14.1	679	1	AL039128	AL039128 DKF2P566K
C 42	462.8	14.0	490	1	AL039589	AL039589 DKF2P434D
C 43	458	13.9	458	6	CA887583	CA887583 B0137H04-
C 44	456.4	13.8	471	1	AL039649	AL039649 DKF2P434G
C 45	449.2	13.6	612	6	CB863409	CB863409 HH04K05W

ALIGNMENTS

RESULT 1
BQ751936
LOCUS BQ751936 DSCT Colletotrichum trifolii cDNA clone pDSC78-67, mRNA
DEFINITION BQ751936 794 bp mRNA linear EST 18-JUL-2002
sequence.
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 794)
AUTHORS Samac D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
Cheung, F. and Fraser, C.M.
ESTS from mycelia of Colletotrichum trifolii race 1
Other ESTs: EST632498
Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbys@puccini.crl.umn.edu
TIGR sequence name: MTSAH67TV More information is available at:
www.medicago.org
Seq primer: lgtA AtA CgA Ctc Act AtA 999 C).
Location/Qualifiers
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/db_xref="taxon:5466"
/clone="pDSC78-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gill from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce.

FEATURES
source

E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

```
Query Match      23.6%; Score 778; DB 5; Length 794;
Best Local Similarity 98.7%; Pred. No. 1.3e-217;
Matches 784; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 GACACATCGAATGGTGCAGAAACCTTTCCGGTATGATAGCGCCGGAAGAGT 62
DB 1 GACACTTTCGAATGGCGCAAAACCTTTCGGTATGATAGCGCCGGAAGAGT 60

QY 63 CAATTCCAGGTGGTGAATGTAACAGTAAAGTTATACGATGTCGAGATATCCCGT 122
DB 61 CAATTCCAGGTGGTGAATGTAACAGTAAAGTTATACGATGTCGAGATATCCCGT 120

QY 123 GTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCTTTCTGCGAAACG 182
DB 121 GTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCTTTCTGCGAAACG 180

QY 183 CCGGAAAAGTGGAAAGCGCGGATGGCGAGCTGAATTAATTCACACCGCGTGSCAAA 242
DB 181 CCGGAAAAGTGGAAAGCGCGGATGGCGAGCTGAATTAATTCACACCGCGTGSCAAA 240

QY 243 CAACTGGCGGGCAACACAGTCGTTGCTGATGCGCGTTGCCACTCCAGTCTGCCCTGAC 302
DB 241 CAACTGGCGGGCAACACAGTCGTTGCTGATGCGCGTTGCCACTCCAGTCTGCCCTGAC 300

QY 303 GCGCGTGCAGAAATGTCGCGCGATTAATCTCCGCGCGATCAACTGGGTGCCAGCGT 362
DB 301 GCGCGTGCAGAAATGTCGCGCGATTAATCTCCGCGCGATCAACTGGGTGCCAGCGT 360

QY 363 GTGGTGTGATGATAGAACGAGCGGCTCGAGCTGTAAAGCGCGGTGCAAACTTT 422
DB 361 GTGGTGTGATGATAGAACGAGCGGCTCGAGCTGTAAAGCGCGGTGCAAACTTT 420

QY 423 CTCGCGCAACCGTCACTGGGCTGATCAATTAATCTCCGCTGGATGACAGGATGCCATT 482
DB 421 CTCGCGCAACCGTCACTGGGCTGATCAATTAATCTCCGCTGGATGACAGGATGCCATT 480

QY 483 GCTGTGGAAGCTGCTGCACTAAATGTTCCGCGGTTATTTCTTGATGTCCTGACAGACA 542
DB 481 GCTGTGGAAGCTGCTGCACTAAATGTTCCGCGGTTATTTCTTGATGTCCTGACAGACA 540

QY 543 CCCATCAACAGTATTATTTCTCCATGAACGAGTACGCGACTGGGCGTGAGCATCTG 602
DB 541 CCCATCAACAGTATTATTTCTCCATGAACGAGTACGCGACTGGGCGTGAGCATCTG 600

QY 603 GTCGATTTGGTCAACGAAATCGCGTGTAGCGGCGCCATTAAAGTTCTGCTCGCGG 662
DB 601 GTCGATTTGGTCAACGAAATCGCGTGTAGCGGCGCCATTAAAGTTCTGCTCGCGG 660

QY 663 CGTCTGCGTCTGGTGGTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
DB 661 CGTCTGCGTCTGGTGGTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720

QY 723 GAAACGGAAGCGAGTGGAGTGCCATGTCGGTTTTCACAAACCATGCAATGCTGAAT 782
DB 721 GAAACGGAAGCGAGTGGAGTGCCATGTCGGTTTTCACAAACCATGCAATGCTGAAT 780

QY 783 GAGGGCATCGTTCC 796
DB 781 GAGGGCATCGTTCC 794
```

RESULT 2
CK394523/c
LOCUS hggada4D08 Gland Cell Amplified cDNA Library Heterodera glycines
DEFINITION

cdna, mRNA sequence.
CK394523
CK394523.1 GI:40389794
EST.
Heterodera glycines
Heterodera glycines
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
1 (bases 1 to 779)
Wang,X., Allen,R., Gao,B., Goellner,M., Maier,T., Baum,T.,
Hussey,R. and Davis,E.
Submission: Wang, X., Allen, R., Gao, B., Goellner, M., Maier, T., Baum, T.,
Hussey, R. and Davis, E.
Unpublished (2003)
Contact: Tom Maier
Department of Plant Pathology, Baum Lab
Iowa State University
351 Bessey Hall, Ames, IA 50011, USA
Tel: 515-294-8854
Fax: 515-294-9420
Email: tmaier@iastate.edu
Heterodera glycines Gland Cell Amplified cDNA Library, single pass
sequence.
Location/Qualifiers
1..779
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell Amplified cDNA Library"
/note="Organ: gland cell; Vector: pSportII"

Query Match 23.4%; Score 770.6; DB 7; Length 779;
Best Local Similarity 99.4%; Pred. No. 2e-215;
Matches 773; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 87 CCAGTAAAGCTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGACCGTTCCCGC 146
DB 779 CCAGTAAAGCTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGACCGTTCCCGC 720
QY 147 GTGCTGAACAGCCAGCCAGCCAGCTTTCTGCGAAAACCGCGGAAAAAGTGAAGCGGCGATG 206
DB 719 GTGCTGAACAGCCAGCCAGCCAGCTTTCTGCGAAAACCGCGGAAAAAGTGAAGCGGCGATG 660
QY 207 GCGGAGCTGAATTAATTCCTCAACCGCGTGGCAACAACAACTGCGCGGCAAAACAGTCGTTG 266
DB 659 GCGGAGCTGAATTAATTCCTCAACCGCGTGGCAACAACAACTGCGCGGCAAAACAGTCGTTG 600
QY 267 CTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCGACGCGCGTCCGAAATTTGTCGGCGG 326
DB 599 CTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCGACGCGCGTCCGAAATTTGTCGGCGG 540
QY 327 ATTAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGTGTCGATGTAGAACGAAGC 386
DB 539 ATTAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGTGTCGATGTAGAACGAAGC 480
QY 387 GCGCTCGAAGCTGTAAAGCGGCGGTGCACATCTTCTCGCCAAACCGGTCTAGTGGGCTG 446
DB 479 GCGCTCGAAGCTGTAAAGCGGCGGTGCACATCTTCTCGCCAAACCGGTCTAGTGGGCTG 420
QY 447 ATCATTAATCTCCGCTGGATGACAGGATGCAATTTGCTGGAAGCTGCCCTGCACTAAT 506
DB 419 ATCATTAATCTCCGCTGGATGACAGGATGCAATTTGCTGGAAGCTGCCCTGCACTAAT 360
QY 507 GTTCGCGGTTATTTCTTGATGTCTGACAGACACCCATCAACAGTATTTTCTCC 566
DB 359 GTTCGCGGTTATTTCTTGATGTCTGACAGACACCCATCAACAGTATTTTCTCC 300
QY 567 CATGAAGACGCTACCGACTGGGCGTGAGCATCTGGTCGATTCGGTCCACGCAATC 626
DB 299 CATGAAGACGCTACCGACTGGGCGTGAGCATCTGGTCGATTCGGTCCACGCAATC 240

QY 627 GCGCTGTTAGCGGCCCATTAAGTTCTCTCGCGCGCTCTGCGTCTGGCTGGTGGCAT 686
|||||
Db 239 GCGCTGTTAGCGGCCCATTAAGTTCTCTCGCGCGCTCTGCGTCTGGCTGGTGGCAT 180
|||||
QY 687 AAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCC 746
|||||
Db 179 AAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCC 120
|||||
QY 747 ATGTCGCGTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTCGCGATG 806
|||||
Db 119 ATGTCGCGTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTCGCGATG 60
|||||
QY 807 CTGTTGCCAAGCATCAGATGGCGTGGCGCAATGCGCGCATTAACGAGTCCGGGC 864
|||||
Db 59 CTGTTGCCAAGCATCAGATGGCGTGGCGCAATGCGCGCATTAACGAGTCCGGGC 2
|||||

RESULT 3
CK118014 769 bp mRNA linear EST 01-JUN-2004
LOCUS 218n06.p1 AtM1 Arabidopsis thaliana cDNA clone MPMGp2011N06218
DEFINITION 5-PRIME, mRNA sequence.

CK118014

CK118014.1 GI:47828330

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 769)

Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.

Generation of a cDNA expression library from Arabidopsis

inflorescence meristem

Unpublished (2003)

Contact: Birgit Kersten

Plant Protein Chip Group, Department Leirach

Max-Planck-Institute for Molecular Genetics

Innestr. 73, D-14195 Berlin, Germany

Tel: +49 (0)30/84131648

Fax: +49 (0)30/84131128

Email: Kersten@molgen.mpg.de

Insert Length: 769 Std Error: 0.00

Plate: 218 row: N column: 6

Seq primer: PQB65.

Location/Qualifiers

1..769

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Columbia"

/db_xref="GABI:956015"

/db_xref="taxon:3702"

/clone="MPMGp2011N06218"

/tissue_type="inflorescence meristem"

/dev_stages="about one week after bolting"

/lab_host="E. coli SCS-1/pSE111"

/clone_lib="ArM1"

/note="Vector: PQB-30NAST-attB (AY386205); Site 1: Sall;

Site 2: NotI; About 1 week after bolting, cDNA synthesis

using SuperscriptTM-system (Invitrogen) with an

oligo (dT)-primer containing NotI restriction site and a

Sall adapter. The main library (plate numbers begin with

1) of 38,000 clones was rearrayed into the sublibrary

(plate numbers begin with 201) containing 5,000 putative

expression clones. Average insert size is 1 kb. Note: The

rearrayed sublibrary (plate numbers begin with 201) was

sequenced. Library generation and sequencing was granted

in context of GABI-LAPP; data are also accessible at

https://gabi.rzpd.de"

ORIGIN

Query Match

Best Local Similarity

23.3%; Score 768; DB 7; Length 769;

100.0%; Pred. No. 1.2e-214;

Matches 768; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 385 GCGCGTCGAAGCCGTGTAAGCGCGGTGCAAAATCTTTCTCGCGCAACGCGTCAGTGGGC 444
Db 1 GCGCGTCGAAGCCGTGTAAGCGCGGTGCAAAATCTTTCTCGCGCAACGCGTCAGTGGGC 60
QY 445 TGATCATTAACATCCGCTGATGACAGAGTGCATTCGTTGTGGAAGCTGCTGCACATA 504
Db 61 TGATCATTAACATCCGCTGATGACAGAGTGCATTCGTTGTGGAAGCTGCTGCACATA 120
QY 505 ATGTTCCGCGCTTAATCTTCTGATGTTCTGACAGACACCCATCAACAGTATTAATTTCT 564
Db 121 ATGTTCCGCGCTTAATCTTCTGATGTTCTGACAGACACCCATCAACAGTATTAATTTCT 180
QY 565 CCCATGAAGACGGTACGCGACTGCGCGTGGAGCATCTGTCGCAATTCGCGTCAACAGCAAA 624
Db 181 CCCATGAAGACGGTACGCGACTGCGCGTGGAGCATCTGTCGCAATTCGCGTCAACAGCAAA 240
QY 625 TCGCGCTGTTAGCGGCCCATTAAGTTCTGTCGCGCGCTCTGCGTCTGGCTGGCTGGC 684
Db 241 TCGCGCTGTTAGCGGCCCATTAAGTTCTGTCGCGCGCTCTGCGTCTGGCTGGCTGGC 300
QY 685 ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTG 744
Db 301 ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTG 360
QY 745 CCATGTCGCGTTTCAACAAACCATGCAAAATTCGTAATGAGGGCATCGTTCCCACTGCGA 804
Db 361 CCATGTCGCGTTTCAACAAACCATGCAAAATTCGTAATGAGGGCATCGTTCCCACTGCGA 420
QY 805 TGCTGGTTGCCAAGCATCAGATGGCGTGGCGCAATGCGCGCATTAACGAGTCCGGGC 864
Db 421 TGCTGGTTGCCAAGCATCAGATGGCGTGGCGCAATGCGCGCATTAACGAGTCCGGGC 480
QY 865 TCGCGCTGTTAGCGGCCCATTAAGTTCTGTCGCGCGCTCTGCGTCTGGCTGGCTGGC 924
Db 481 TCGCGCTGTTAGCGGCCCATTAAGTTCTGTCGCGCGCTCTGCGTCTGGCTGGCTGGC 540
QY 925 ATATCCCGCGTTAAACCAACCATGCAAAATTCGCTGCTGGGGCAACACGCGTGG 984
Db 541 ATATCCCGCGTTAAACCAACCATGCAAAATTCGCTGCTGGGGCAACACGCGTGG 600
QY 985 ACCGTTGCTGCAACTCTCTCAGGGCCAGGGCGGTGAAGGGCAATCAGCTGTTCGCCGTCT 1044
Db 601 ACCGTTGCTGCAACTCTCTCAGGGCCAGGGCGGTGAAGGGCAATCAGCTGTTCGCCGTCT 660
QY 1045 CACTGGTGAAAAGAAAACCAACCCCTGGCGCCCAATACGCAAAACCGCTCTCCCCCGCGCT 1104
Db 661 CACTGGTGAAAAGAAAACCAACCCCTGGCGCCCAATACGCAAAACCGCTCTCCCCCGCGCT 720
QY 1105 TGGCCGATTTCATTAATGAGCTGGCAGCAGAGTTTCCCGACTGGAAA 1152
Db 721 TGGCCGATTTCATTAATGAGCTGGCAGCAGAGTTTCCCGACTGGAAA 768
RESULT 4
BQ751655
LOCUS BQ751655
DEFINITION EST632218 DSCT Colletotrichum trifolii cDNA clone pDSCT6-59, mRNA
sequence.
ACCESSION BQ751655
VERSION BQ751655.1 GI:21907060
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 793)
AUTHORS Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
Cheung,F. and Fraser,C.M.
TITLE ESTs from mycelia of Colletotrichum trifolii race 1
JOURNAL Unpublished (2002)

COMMENT

Other ESTs: EST632217
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org
 Seq primer: (gta ata cga ctc act ata ggg c).

FEATURES

source

Location/Qualifiers

1..793

/organism="Colletotrichum trifolii"

/mol_type="mRNA"

/strain="race 1"

/db_xref="taxon:5466"

/clone="pDST6-59"

/tissue_type="mycelia"

/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."

/lab_host="DH5alpha"

/clone_lib="DST6"

/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ enriched RNA. The cDNA was ligated into Lambda gtl1 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 23.1%; Score 762; DB 5; Length 793;
 Best Local Similarity 99.1%; Pred. No. 7e-213;
 Matches 777; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 3 GACACCATCGAATGGTGCAAAACCTTTCGCGGTATGCGATGATGCGAGAGATGCCGGT 122
 DB 1 GACATCTCGAATGGCGCAAAACCTTTCGCGGTATGCGATGATGCGAGAGATGCCGGT 60

QY 63 CAATTCAGGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTCGAGAGATGCCGGT 122
 DB 61 CAATTCAGGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTCGAGAGATGCCGGT 120

QY 123 GTCTCTTATCAGACCGTTTCCCGGCGGTGTAACAGCCAGCCAGCTTTCTGCGAAAACG 182
 DB 121 GTCTCTTATCAGACCGTTTCCCGGCGGTGTAACAGCCAGCCAGCTTTCTGCGAAAACG 180

QY 183 CGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTAACATTCCTCAACCGCGTGGACAA 242
 DB 181 CGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTAACATTCCTCAACCGCGTGGACAA 240

QY 243 CAATCGCGGGCAACAGTGTGCTGATTTGGCGGTGTCACCTCCAGTTCGGCCCTGCAC 302
 DB 241 CAATCGCGGGCAACAGTGTGCTGATTTGGCGGTGTCACCTCCAGTTCGGCCCTGCAC 300

QY 303 CGCGCGTCGCAAAATTTGTCGGCGGATTAATTCGCGCGGATCACTGCGTGGCGCGGTG 362
 DB 301 CGCGCGTCGCAAAATTTGTCGGCGGATTAATTCGCGCGGATCACTGCGTGGCGCGGTG 360

QY 363 GTGGTGTGATGTAGAACGAAGCGGCTCGAAGCCCTGTAAAGCGGCGGTGCAACAATCTT 422
 DB 361 GTGGTGTGATGTAGAACGAAGCGGCTCGAAGCCCTGTAAAGCGGCGGTGCAACAATCTT 420

QY 423 CTCGCGCAACCGCTCAGTGGCGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 482
 DB 421 CTCGCGCAACCGCTCAGTGGCGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 480

QY 483 GCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACGACA 542
 DB 481 GCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACGACA 540

QY 543 CCCATCAACAGTATATTTCTCCCATGAGACGGTAGCGACTGGCGGTGGAGATCTG 602
 DB 541 CCCATCAACAGTATATTTCTCCCATGAGACGGTAGCGACTGGCGGTGGAGATCTG 600

QY 603 GTCGATTTGGTTCACCAAGCAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGGCG 662
 DB 601 GTCGATTTGGTTCACCAAGCAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGGCG 660

QY 663 CGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
 DB 661 CGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720

QY 723 GAACGGGAAGCGCACTGGAGTGGCATGTCGGTTCCTCAACCAACCATGCAAAATGCTGAAT 782
 DB 721 GAACGGGAAGCGCACTGGAGTGGCATGTCGGTTCCTCAACCAACCATGCAAAATGCTGAAT 778

QY 783 GAGG 786
 DB 779 GAGG 782

RESULT 5

BO751220

LOCUS

DEFINITION

BO751220

ACCESSION

BO751220.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..806

/organism="Colletotrichum trifolii"

/mol_type="mRNA"

/strain="race 1"

/db_xref="taxon:5466"

/clone="pDST3-64"

/tissue_type="mycelia"

/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."

/lab_host="DH5alpha"

/clone_lib="DST6"

/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ enriched RNA. The cDNA was ligated into Lambda gtl1 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid

TIGR sequence name: MTSAC64TV More information is available at:

www.medicago.org

Seq primer: (gta ata cga ctc act ata ggg c).

lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN		Query Match		22.9%; Score 755.4; DB 5; Length 806;		Best Local Similarity		97.4%; Pred. No. 6.3e-211;		Matches 790; Conservative 0; Mismatches 16; Indels 5; Gaps 2;	
Qy	3	GACACCATCGAATGTGTGCAAAACCTTTTCGGTATGATGATAGCGCCGGAAGAGAGT	62			1	GACATTTGCAATGGCGCAAAACCTTTTCGGTATGATGATAGCGCCGGAAGAGAGT	60			
Db	1	GACATTTGCAATGGCGCAAAACCTTTTCGGTATGATGATAGCGCCGGAAGAGAGT	60			Qy	63	CAATTCAGGGTGGTGAATGTGAAACACAGTATACGATGTCGACAGATATGCCGGT	122		
Db	61	CAATTCAGGGTGGTGAATGTGAAACACAGTATACGATGTCGACAGATATGCCGGT	120			Db	61	CAATTCAGGGTGGTGAATGTGAAACACAGTATACGATGTCGACAGATATGCCGGT	120		
Qy	123	GTCTCTATCAGACGTTTCGGCGTGTGAACACAGGCGACGACGATTTCTGCGAAACG	182			Db	121	GTCTCTATCAGACGTTTCGGCGTGTGAACACAGGCGACGACGATTTCTGCGAAACG	180		
Qy	183	CGGGAAGAAGTGGAGCGCGATGCGGAGCTGAATTAACATTCACATTCACACCGGTGCAAA	242			Db	181	CGGGAAGAAGTGGAGCGCGATGCGGAGCTGAATTAACATTCACATTCACACCGGTGCAAA	240		
Qy	243	CAACTGGCGGCAACACAGTGTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT	302			Db	241	CAACTGGCGGCAACACAGTGTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT	300		
Qy	303	CGCGCGTGCAGAAATGTTCGGCGGATTAATATCTCGCGCGATCAACTGGGTGCCAGCGTG	362			Db	301	CGCGCGTGCAGAAATGTTCGGCGGATTAATATCTCGCGCGATCAACTGGGTGCCAGCGTG	360		
Qy	363	GTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	422			Db	361	GTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	420		
Qy	423	CTCGCGCAACGCTCAGTGGGCTGATCAATTAACATTCGCTGGATGACAGGATGCCATT	482			Db	421	CTCGCGCAACGCTCAGTGGGCTGATCAATTAACATTCGCTGGATGACAGGATGCCATT	480		
Qy	483	GCTGTGGAAGTGTGCTGACATTAATGTTCCGGGTTATTTCTTGATGTCTTGACACAGACA	542			Db	481	GCTGTGGAAGTGTGCTGACATTAATGTTCCGGGTTATTTCTTGATGTCTTGACACAGACA	540		
Qy	543	CCCATCAACAGTATTTTCTCCCATGACGCGTACCGACTGCGGCTGGGCTGGAGCATCTG	602			Db	541	CCCATCAACAGTATTTTCTCCCATGACGCGTACCGACTGCGGCTGGGCTGGAGCATCTG	600		
Qy	603	GTGCGATTGGGTACACAGCAATCGCGTGTAGCGGGCCCAATTAAGTTCTGCTCGGCG	662			Db	601	GTGCGATTGGGTACACAGCAATCGCGTGTAGCGGGCCCAATTAAGTTCTGCTCGGCG	660		
Qy	663	CGTCTGCGTCTGGCTGGCGATTAATATCTCACTCGCAATCAAATTCAGCCGATAGCG	722			Db	661	CGTCTGCGTCTGGCTGGCGATTAATATCTCACTCGCAATCAAATTCAGCCGATAGCG	720		
Qy	723	GAAACGGGAAGCGACTGAGTGCATATGTCGGGTTTTCACAAACCAACATGCAATGCTGAAT	782			Db	721	GAAACGGGAAGCGACTGAGTGCATATGTCGGGTTTTCACAAACCAACATGCAATGCTGAAT	782		
Qy	783	GAGGCGATCGTTCCACATCGCATGCTGGTTG	813			Db	778	GAGGCGATCGTTCCACATCGCATGCTGGTTG	806		

RESULT 6
CB863541/c
LOCUS HH04A08y HH Hordeum vulgare cDNA clone HH04A08 3-PRIME, mRNA
DEFINITION CB863541 689 bp mRNA linear EST 22-APR-2003
sequence.

CB863541
CB863541.1 GI:30058100
EST.
Hordeum vulgare
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 689)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
Graner,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 689 Std Error: 0.00
Plate: 4 row: A column: 8
Seq primer: SP6.
Location/Qualifiers
1..689
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555198"
/db_xref="taxon:4513"
/clone="HH04A08"
/tissue_type="coleoptile"
/dev_stage="coleoptile, 1 day old"
/lab_host="DH10B"
/clone_lib="HH"
/note="Vector: pSPORT; Site 1: SalI (5-end of cDNA);
Site 2: NciI (3-end of cDNA); Due to the cloning system
used blue/white selection for recombinants is not 100 %
reliable. Average insert size is 1.3 Kb."

ORIGIN		Query Match		20.8%; Score 684.8; DB 6; Length 689;		Best Local Similarity		99.6%; Pred. No. 4.2e-190;		Matches 686; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Qy	21	AAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGAGTCAATTCAGGTGGTGAAT	80			Db	689	AAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGAGTCAATTCAGGTGGTGAAT	630		
Qy	81	GTGAAACCAAGTAAACGTTATAGATGTCGACAGTATGCGGTGTCTCTTATCAGACCGTT	140			Db	629	GTGAAACCAAGTAAACGTTATAGATGTCGACAGTATGCGGTGTCTCTTATCAGACCGTT	570		
Qy	141	TCCCGCGTGGTGAACCGACCGCCAGCTTTCTCGAAGAACCGCGGAAAGTGGAAAGC	200			Db	569	TCCCGCGTGGTGAACCGACCGCCAGCTTTCTCGAAGAACCGCGGAAAGTGGAAAGC	510		
Qy	201	GCATGGCGGAGCTGAATTAATTCACATTCACCAACCGCTGGCACAACAACTGGCGGCAACAG	260			Db	509	GCATGGCGGAGCTGAATTAATTCACATTCACCAACCGCTGGCACAACAACTGGCGGCAACAG	450		
Qy	261	TCGTTGCTGATTGGCGTTGCCACCTTCCAGTCTGGCCCTTCACCGCCGCTCGCAAAATTGTC	320			Db	449	TCGTTGCTGATTGGCGTTGCCACCTTCCAGTCTGGCCCTTCACCGCCGCTCGCAAAATTGTC	390		
Qy	321	CGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGGTGGTGTGATGATGATGATGAT	380			Db	389	CGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGGTGGTGTGATGATGATGATGAT	330		
Qy	381	CGAAGCGCGTGAAGAGCTGTAAAGCGCGGTGCAATCTTCTCGCGCAACCGGTGATGAT	440			Db	329	CGAAGCGCGTGAAGAGCTGTAAAGCGCGGTGCAATCTTCTCGCGCAACCGGTGATGAT	270		

```

QY 441 GGCTGATCAATTAACATATCCGCTGGATGACAGATGCCATTGCTGTGGAAGCTGCCTGC 500
Db 269 GGCTGATCAATTAACATATCCGCTGGATGACAGATGCCATTGCTGTGGAAGCTGCCTGC 210
QY 501 ACTAATGTTCCGGCGTATTCTTGTATGCTCTGACAGACACCCATCAACAGTATTATT 560
Db 209 ACTAATGTTCCGGCGTATTCTTGTATGCTCTGACAGACACCCATCAACAGTATTATT 150
QY 561 TTCTCCCATGAAGACGGTACGGACACTGGCGGTGGAGCATCTGCTGCAATGGGTGACAG 620
Db 149 TTCTCCCATGAAGACGGTACGGACACTGGCGGTGGAGCATCTGCTGCAATGGGTGACAG 90
QY 621 CAAATCGCCTGTTAGCGGCGCCATTAAGTTCTGCTCGGCGCTGCTGCTGCTGCTGCTG 680
Db 89 CAAATCGCCTGTTAGCGGCGCCATTAAGTTCTGCTCGGCGCTGCTGCTGCTGCTGCTG 30
QY 681 TGGCATAAATATCTCACTCGCAATCAAT 709
Db 29 TGGCATAAATATCTCACTCGCAATCAAT 1

RESULT 7
AL038548/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL038548 712 bp mRNA linear EST 04-SEP-2003
DKFZ566E0346_r1 566 (synonym: hfkd2) Homo sapiens cDNA clone
DKFZ566E0346 5', mRNA sequence.
AL038548
AL038548.1 GI:5407738
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Ottenwaelder, et al.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZ566E0346) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..712
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZ566E0346"
/tissue_type="kidney"
/dev_stage="fetal"
/lab_host="X1-2blue"
/clone_lib="566 (synonym: hfkd2)"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

Query Match
Best Local Similarity 20.5%; Score 676.4; DB 1; Length 712;
Matches 699; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 33 GGTATGCGATGATAGCCCGGAGAGAGTCAATTGAGGTGGTGAATGGAACAGTA 92
Db 712 GGTATGCGATGATAGCCCGGAGAGAGTCAATTGAGGTGGTGAATGGAACAGTA 653
QY 93 ACCTTATACATGTCGAGAGTATGCCGGTGTCTTTATCAGACCGTTCCCGCGGTG 152

```

```

Db 652 ACCTTATACATGTCGAGAGTATGCCGGTGTCTTTATCAGACCGTTTCCCGGTGGT 593
QY 153 AACCAGGCCAGCCAGTTTCTGCGAAACCGCGG-AAAAAGTGAAGCGCGGATGGCGGA 211
Db 592 AACCAGGCCAGCCAGTTTCTGCGAAACCGCGGAAAAAGTGAAGCGCGGATGGCGGA 533
QY 212 GCTGAATTAACATTCGCCAACCGCGTGGGCAACAACCTGGCGGGCAACAGTGTGCTGAT 271
Db 532 GCTGAATTAACATTCGCCAACCGCGTGGGCAACAACCTGGCGGGCAACAGTGTGCTGAT 473
QY 272 TGGCGTTGCCACCTCCAGTCTGGCCCTGCGAC-GGCGCGTGCAGAAATTTGTCGGCGGATTA 330
Db 472 TGGCGTTGCCACCTCCAGTCTGGCCCTGCGACGGCGCGCTGCAGAAATTTGTCGGCGGATTA 413
QY 331 AATCTCGCGCGGATCAACTGGGTGCCAGCGTGGTGTGTCGATGCTAGAAACGAAGCGCG 390
Db 412 AATCTCGCGCGGATCAACTGGGTGCCAGCGTGGTGTGTCGATGCTAGAAACGAAGCGCG 353
QY 391 TCGAAGCGCTGTAAAGCGCGGTGCACAAATCTTCTCGCGCAACCGGTGAGTCA 450
Db 352 TCGAAGCGCTGTAAAGCGCGGTGCACAAATCTTCTCGCGCAACCGGTGAGTCA 293
QY 451 TTAACATATCCGCTGGATGACAGATGCCATTGCTGTGGAAGCTGCCTGCACATAATGTT 510
Db 292 TTAACATATCCGCTGGATGACAGATGCCATTGCTGTGGAAGCTGCCTGCACATAATGTT 233
QY 511 CGCGGTATTCTTGTATGCTCTGACAGACACCCATCAACAGTATTATTTCTCCCATG 570
Db 232 CGCGGTATTCTTGTATGCTCTGACAGACACCCATCAACAGTATTATTTCTCCCATG 173
QY 571 AAGACGGTACCGCACTGGCGTGGGATGCGCATCTGTCGCATTTGGTCCAGCAATCGCG 630
Db 172 AAGACGGTACCGCACTGGCGTGGGATGCGCATCTGTCGCATTTGGTCCAGCAATCGCG 113
QY 631 TGTAGCGGCGCCATTAAGTTCTGTCGCGCGCTGTCGCTGCTGGTGGTGGCATAAAT 690
Db 112 TGTAGCGGCGCCATTAAGTTCTGTCGCGCGCTGTCGCTGCTGGTGGTGGCATAAAT 53
QY 691 ATCTCACTCGCAATCAATTCAGCCCATAGCGGAACGGGAAG 732
Db 52 ATCTCACTCGCAATCAATTCAGCCCATAGCGGAACGGGAAG 11

AV080106 710 bp DNA linear GSS 06-NOV-2002
AY080106 Scripps Pier (La Jolla, CA) uncultured virus community
uncultured marine virus genomic clone SIO51p3G6L, genomic survey
sequence.
AY080106
AY080106.1 GI:24745302
GSS.
uncultured marine virus
uncultured marine virus
Viruses; environmental samples.
1 (bases 1 to 710)
Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
Mead,D., Azam,F. and Rohwer,F.
Genomic analysis of uncultured marine viral communities
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
22294988
12384570
Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@sunstroke.sdsu.edu
Class: shotgun.
Location/Qualifiers
1..710
/organism="uncultured marine virus"

```

```
/mol_type="genomic DNA"  
/db_xref="taxon:186617"  
/clone="SI05lpj36L"  
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus  
community"  
/note="Marine viruses were isolated from 200 liters of  
surface seawater using a combination of differential  
filtration and density-dependent gradient centrifugation.  
Linker-amplified shotgun libraries were constructed by  
randomly shearing the total marine viral community DNA,  
end-repairing, ligating dsDNA linkers to the ends, and  
amplifying the fragments using Vent DNA polymerase. The  
resulting fragments were ligated into the pSMART vector  
and electroporated into MC12 cells (Lucigen; Middleton,  
WI)."
```

ORIGIN

```
Query Match      20.5%; Score 676; DB 8; Length 710;  
Best Local Similarity 98.9%; Pred. No. 1.7e-187;  
Matches 690; Conservative 0; Mismatches 7; Indels 1; Gaps 1;  
  
QY 400 GTAAAGCGCGTGCACAACTTCTCGCGCAACGGCTCAGTGGGCTGATCACTAACTATC 459  
DB |||||  
8 GTGCTGCAGCGTGCACAACTTCTCGCGCAACGGCTCAGTGGGCTGATCACTAACTATC 67  
  
QY 460 CGCTGGATGACAGATGCGCATGCTGTGGAAAGCTGCGTGCATTAATGTTCCGGCGTTAT 519  
DB |||||  
68 CGCTGGATGACAGATGCGCATGCTGTGGAAAGCTGCGTGCATTAATGTTCCGGCGTTAT 127  
  
QY 520 TTCTTGATGTCTGCACGACACCATCAACAGATATTATTTCTCCCATGAAGACGGTA 579  
DB |||||  
128 TTCTTGATGTCTGCACGACACCATCAACAGATATTATTTCTCCCATGAAGACGGTA 187  
  
QY 580 CGCGACTGGCGTGAGCATCTGTGCGCATGCGGTCAACGACAAATCGCGCTGTTAGCGG 639  
DB |||||  
188 CGCGACTGGCGTGAGCATCTGTGCGCATGCGGTCAACGACAAATCGCGCTGTTAGCGG 247  
  
QY 640 GCCCATTAAGTTCTGTCTCGGGCGTCTGCTGCTGGCTGGCTGGCATAAATATCTCACTC 699  
DB |||||  
248 GCCCATTAAGTTCTGTCTCGGGCGTCTGCTGCTGGCTGGCTGGCATAAATATCTCACTC 307  
  
QY 700 GCAATCAATTCAGCCGATAGCGGAACGGAAGGCGACTGGAGTGCCATGTCGGGTTTC 759  
DB |||||  
308 GCAATCAATTCAGCCGATAGCGGAACGGAAGGCGACTGGAGTGCCATGTCGGGTTTC 367  
  
QY 760 AACAAACCATGCAATGCTGAATGAGGCGATCGTTCCCACTCGCATGCTGGTGGCAACG 819  
DB |||||  
368 AACAAACCATGCAATGCTGAATGAGGCGATCGTTCCCACTCGCATGCTGGTGGCAACG 427  
  
QY 820 ATCAGATGCGCTGGCGCAATGCGCGCATTTACCGAGTCCGGGCTGCGCGTGGTGGCGG 879  
DB |||||  
428 ATCAGATGCGCTGGCGCAATGCGCGCATTTACCGAGTCCGGGCTGCGCGTGGTGGCGG 487  
  
QY 880 ATATCTCGGTAGTGGATACGACATACCGAAGACAGCTCATGTTATATCCCGCGTTAA 939  
DB |||||  
488 ATATCTCGGTAGTGGATACGACATACCGAAGACAGCTCATGTTATATCCCGCGTTAA 547  
  
QY 940 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACAGCGTGGACCGCTTGTGCGCAAC 999  
DB |||||  
548 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACAGCGTGGACCGCTTGTGCGCAAC 607  
  
QY 1000 TCTCTCAGGGCGAGCGGTGAAGGGCAATCAGCTGTTGCGCTCTCACTGGTGAAGAA 1059  
DB |||||  
608 TCTCTCAGGGCGAGCGGTGAAGGGCAATCAGCTGTTGCGCTCTCACTGGTGAAGAA-AA 666  
  
QY 1060 AAACCACTCGGGCCCATAGCGAACCGCTCTCCC 1097  
DB |||||  
667 AAACCACTCGGGCCCATAGCGAACCGCTCTCCC 704
```

RESULT 9

AL044483/c

LOCUS

660 bp mRNA linear EST 04-SEP-2003

AL044483

```
DEFINITION DKFZp4341102_s1 434 (synonym: htes3) Homo sapiens cDNA clone  
ACCESSION DKFZp4341102.3, mRNA sequence.  
VERSION AL044483  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 660)  
AUTHORS Ansorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and  
Wiemann,S.  
TITLE EST (Ansorge, Benes, et al.)  
JOURNAL Unpublished (1999).  
COMMENT Contact: MIPS
```

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 3' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by EMBL (European Molecular Biology Laboratories,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No r1 sequence available.

This clone (DKFZp4341102) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

FEATURES

source

```
1..660  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="DKFZp4341102"  
/tissue_type="testis"  
/dev_stage="adult"  
/lab_host="DH10B"  
/clone_lib="434 (synonym: htes3)"  
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
```

ORIGIN

```
Query Match      19.6%; Score 647.4; DB 1; Length 660;  
Best Local Similarity 99.7%; Pred. No. 4.5e-179;  
Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
  
QY 85 AACCAAGTAACTATACGATGTCCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCC 144  
DB |||||  
660 AACCAAGTAACTATACGATGTCCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCC 601  
  
QY 145 GGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAACCGGGGAAAGTGGAGCGGCA 204  
DB |||||  
600 GGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAACCGGGGAAAGTGGAGCGGCA 541  
  
QY 205 TGGCGGAGCTGAATTAACATTTCCCAACCGTGGGCAACAACCTGGCGGCAACAGTCGT 264  
DB |||||  
540 TGGCGGAGCTGAATTAACATTTCCCAACCGTGGGCAACAACCTGGCGGCAACAGTCGT 481  
  
QY 265 TGCTGATTGGCGTGGCCACCTCCAGTCTGGCCCTGCACGCGCGCTCGCAAAATTTGCGCG 324  
DB |||||  
480 TGCTGATTGGCGTGGCCACCTCCAGTCTGGCCCTGCACGCGCGCTCGCAAAATTTGCGCG 421  
  
QY 325 CGATTAAATCTCGCGCGCATCACTGGGTGCAGCGTGGTGTGTCGATGGTAGAACGAA 384  
DB |||||  
420 CGATTAAATCTCGCGCGCATCACTGGGTGCAGCGTGGTGTGTCGATGGTAGAACGAA 361  
  
QY 385 GCGGCGTGAACCGCTGTAAGCGGCGTGCAAAATCTTCTCGCGCAACCGCTCAGTGGGC 444  
DB |||||  
360 GCGGCGTGAACCGCTGTAAGCGGCGTGCAAAATCTTCTCGCGCAACCGCTCAGTGGGC 301  
  
QY 445 TGATCATTAATCATCCGCTGGATGACCAAGATGCCATTCCTGTGGAAGCTGCTGCACCTA 504  
DB |||||  
300 TGATCATTAATCATCCGCTGGATGACCAAGATGCCATTCCTGTGGAAGCTGCTGCACCTA 241  
  
QY 505 ATGTTCCGGCGTTATTTCTTGATGTCTCTGACGACAGACCCCATCAACAGTATATTTTCT 564
```

```

Db      240  ATGTTCCGGGTTATTTCCTTGATGTCCTGACCAGACCCCATCAACAGTATTATTTCT 181
QY      565  CCCATGAAGACGGTAGCGACTGGGCGTGGAGCATCTGTCGATTTGGGTACACGAGAAA 624
Db      180  CCATGAAGACGGTAGCGACTGGGCGTGGAGCATCTGTCGATTTGGGTACACGAGAAA 121
QY      625  TCGCGCTGTTAGCGGCCCAATTAAGTTCTCTCGGCGGCTCTGGGTCTGGCTGGCTGGC 684
Db      120  TCGCGCTGTTAGCGGCCCAATTAAGTTCTCTCGG- GCGTCTGGCTGGCTGGCTGGC 62
QY      685  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAAACGGGAAGCGACTGGAGTG 744
Db      61  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAAACGGGAAGCGACTGGAGTG 2
QY      745  C 745
Db      1  C 1

RESULT 10
BJ063872/c
LOCUS   BJ063872 NTBB Mochii normalized Xenopus tailbud library Xenopus
DEFINITION
ACCESSION BJ063872
VERSION   BJ063872.1 GI:17471074
KEYWORDS EST.
SOURCE    Xenopus laevis (African clawed frog)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
          Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 641)
AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
          Kohara,Y.
TITLE     Expressed genes in X. laevis embryo
JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-i
          Center For Genetic Resource Information
          National Institute of Genetics
          1111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tsunigenes.nig.ac.jp
          The information of this clone is available through the following
          URL.
          http://xenopus.nibb.ac.jp.

FEATURES             source
     Location/Qualifiers
         1..641
            /organism="Xenopus laevis"
            /mol_type="mRNA"
            /db_xref="taxon:8355"
            /clone="XL077k02"
            /tissue_type="whole embryo"
            /dev_stage="stage 25"
            /clone_lib="NIBB Mochii normalized Xenopus tailbud
            library"

ORIGIN
Query Match      19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred. No. 3.5e-177;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30  CGCGGTATGGCATGATAGCGCCCGAAGAGAGTCAATTCAGGTGGTGAATGTGAACCA 89
Db      641  CGCGGTATGGCATGATAGCGCCCGAAGAGAGTCAATTCAGGTGGTGAATGTGAACCA 582
QY      90  GTAACGTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCGGGTG 149
Db      581  GTAACGTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCGGGTG 522
QY      150  GTGAACACGAGCCAGCCACGTTTCTCGGAAACCGGGGAAAAAGTGGGAAGCGCGGATGGCG 209

```

```

Db      521  GTGAACACGAGCGAGCCAGTTTCTGCGAAAAACGGGGAAAAAGTGAAGCGCGATGGCG 462
QY      210  GAGCTGAATTACATTTCCCAACCGGTGGCAACAACACTGGCGGGCAACAGTTCGTTGCTG 269
Db      461  GAGCTGAATTACATTTCCCAACCGGTGGCAACAACACTGGCGGGCAACAGTTCGTTGCTG 402
QY      270  ATTGGCGTTGCCACTCCAGTCTGGCCCTGCAACGCGCGTGCAGAAATTTGTCGCGCGGAT 329
Db      401  ATTGGCGTTGCCACTCCAGTCTGGCCCTGCAACGCGCGTGCAGAAATTTGTCGCGCGGAT 342
QY      330  AAATCTCGCGCGGATCAACTGGGTGCCAGCGTGGTGTGTCGATGGTAGAACGAAGCGGC 389
Db      341  AAATCTCGCGCGGATCAACTGGGTGCCAGCGTGGTGTGTCGATGGTAGAACGAAGCGGC 282
QY      390  GTCGAAGCCTGTAAAGCGGGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 449
Db      281  GTCGAAGCCTGTAAAGCGGGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 222
QY      450  ATTAACATCTCGCTGGATGACAGGATGCCATGCTGTGGAAGCTGCTGCACCTAAATGTT 509
Db      221  ATTAACATCTCGCTGGATGACAGGATGCCATGCTGTGGAAGCTGCTGCACCTAAATGTT 162
QY      510  CCGGCGTTATTTCTTGATGTCCTGACCAACACCCATCAACAGTATTATTTCTCCCAT 569
Db      161  CCGGCGTTATTTCTTGATGTCCTGACCAACACCCATCAACAGTATTATTTCTCCCAT 102
QY      570  GAAGACGGTACGGACTGGCGTGGAGCATCTGTCGCAATGGGTCAACAGCAATCGCG 629
Db      101  GAAGACGGTACGGACTGGCGTGGAGCATCTGTCGCAATGGGTCAACAGCAATCGCG 42
QY      630  CTGTTAGCGGGCCCATTAAGTTCTGTCGCGCGGCTGGCG 670
Db      41  CTGTTAGCGGGCCCATTAAGTTCTGTCGCGCGGCTGGCG 1

RESULT 11
CB863814/c
LOCUS   CB863814 HH Hordeum vulgare cDNA clone HH07B20 3-PRIME, mRNA
DEFINITION
ACCESSION CB863814
VERSION   CB863814.1 GI:30058373
KEYWORDS EST.
SOURCE    Hordeum vulgare
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
          Poideae; Triticeae; Hordeum.
          Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
          Graner,A.
          Barley ESTs from coleoptile tissue
          Unpublished (2003)
          Contact: Stein Nils
          Molecular Markers Group, Department Genbank
          Institute of Plant Genetics and Crop Plant Research (IPK)
          Corrensstr. 3, 06466, Gatersleben, Germany
          Tel: 039482-5522
          Fax: 039482-5595
          Email: stein@ipk-gatersleben.de
          Insert Length: 624 Std Error: 0.00
          Plate: 7 row: B column: 20
          Seq primer: SP6.
          Location/Qualifiers
              1..624
                 /organism="Hordeum vulgare"
                 /mol_type="mRNA"
                 /culturivar="Sloop"
                 /db_xref="GABI:555471"
                 /db_xref="taxon:4513"
                 /clone="HH07B20"
                 /tissue_type="coleoptile"

```


EST (Duesterhoeft, et al.)		Unpublished (1999)		Contact: MIPS		Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany		This is the 3' sequence of the clone insert		Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); sequenced by Qiagen within the cDNA sequencing consortium of the German Genome Project. r1 sequence also available.		This clone is available at the RZPD in Berlin.		Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.		Location/Qualifiers																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
1..794		/organism="Homo sapiens"		/mol_type="mRNA"		/db_xref="taxon:9606"		/clone="DKFZp434B075"		/tissue_type="testis"		/dev_stage="adult"		/lab_host="DH10B"		/clone_lib="434 (synonym: htes3)"		/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
FEATURES		source		18.5%; Score 612; DB 1; Length 794;		Best Local Similarity 99.7%; Pred. No. 1.4e-168;		Matches 634; Conservative 0; Mismatches 0; Indels 2; Gaps 2;		569		TGAAGACGGTAGCGACTGGGCGTGGAGCATCTGTCGCATTTGGTCCAGGCAATCGC		628		794		TGAAGACGGTAGCGACTGGGCGTGGAG-ATCTGTCGCATTTGGGTCCAGCAATCGC		736																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY		Db		QY	

```

RESULT 13
CL655575/c
LOCUS
DEFINITION
PR10124a_C01 - PRI0124a.B21 (752) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL655575
CL655575.1 GI:50134897
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 752)
AppADB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: raif.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
Location/Qualifiers
1..752
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/notes="Vector: pBpifos-5 Fosmid vector"

ORIGIN
Query Match 18.3%; Score 603; DB 9; Length 752;
Best Local Similarity 98.9%; Pred. No. 6e-166;
Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Qy 1996 TTCACCTGGCCGCTGATTCGCTGAC -GGGGGTTATGCGTTCAAGTATGAAACGGCAA 2054
Db 752 TTCACCTGGCCGCTGATTCGCTGACGGGGGTTATGCGTTCAAGTATGAAACGGCAA 693

Qy 2055 GTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAACGGGCTCTGACCTTCCT 2114
Db 692 GTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAACGGGCTCTGACCTTCCT 633

Qy 2115 GGTTCACCTGATTAAGCAACACATGATGCGACACCGATTCCTCCATCGCAGAAGC 2174
Db 632 GGTTCACCTGATTAAGCAACACATGATGCGACACCGATTCCTCCATCGCAGAAGC 573

Qy 2175 TGCCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCGCTGGGCATGCTCCACAT 2234
Db 572 TGCCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCGCTGGGCATGCTCCACAT 513

Qy 2235 CGACACAGCAAGTGAATTAATGTTATGCGTAAACGCTGCGACCTTCAAGGGTCAACCATC 2294
Db 512 CGACACAGCAAGTGAATTAATGTTATGCGTAAACGCTGCGACCTTCAAGGGTCAACCATC 453

Qy 2295 CAACCGCTGCTTGGCGTGTGAGCGAGGTATTAACGCCCGCAGTCCGACAAAGGCT 2354
Db 452 CAACCGCTGCTTGGCGTGTGAGCGAGGTATTAACGCCCGCAGTCCGACAAAGGCT 393

Qy 2355 GGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGGCTGGAACGGGTAA-TA 2413
Db 392 GGGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGGCTGGAACGGGTAA-TA 333

```

```

Qy 2414 AAGACAAACCGCTGGGTGCGGTAGCTTACGAGGAAGAGTTGCGGAAAGATC 2473
Db 332 AGGCCAACCCGCTTGGGTGCGGTAGCTTACGAGGAAGAGTTGCGGAAAGATC 273

Qy 2474 CACGTATTGCGCGCCACCATGGAAAACCCCGAAGGTGAATCATGCCGAACATCCCGC 2533
Db 272 CACGTATTGCGCGCCACCATGGAAAACCCCGAAGGTGAATCATGCCGAACATCCCGC 213

Qy 2534 AGATGTCGCGCTTTCTGTTATGCGCTGCTACTGCGTGATCAACGCCGCCGCGGTGTC 2593
Db 212 AGATGTCGCGCTTTCTGTTATGCGCTGCTACTGCGTGATCAACGCCGCCGCGGTGTC 153

Qy 2594 AGACTGTCGATGAAGCCCTGAACACGCGCAGACT 2628
Db 152 AGACTGTCGATGAAGCCCTGAACACGCGCAGACT 118

RESULT 14
AV080094/c
LOCUS
DEFINITION
AY080094 Scripps Pier (La Jolla, CA) uncultured virus community
sequence.
AY080094
AY080094.1 GI:24745276
GSS.
uncultured marine virus
uncultured marine virus
Viruses; environmental samples.
1 (bases 1 to 608)
Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M.,
Mead, D., Azam, F. and Rohwer, F.
Genomic analysis of uncultured marine viral communities
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
22294988
PUBMED
12384570
COMMENT
Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@sunstroke.sdsu.edu
Class: shotgun.
Location/Qualifiers
1..608
/organism="uncultured marine virus"
/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone_lib="SIO51p3D4L"
/community="Marine viruses were isolated from 200 liters of
surface seawater using a combination of differential
filtration and density-dependent gradient centrifugation.
Linker-amplified shotgun libraries were constructed by
randomly shearing the total marine viral community DNA,
end-repairing, ligating dEDNA linkers to the ends, and
amplifying the fragments using Vent DNA polymerase. The
resulting fragments were ligated into the pSMART vector
and electroporated into MC12 cells (Lucigen; Middleton,
WI)"

ORIGIN
Query Match 17.9%; Score 589.8; DB 8; Length 608;
Best Local Similarity 99.7%; Pred. No. 4.5e-162;
Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 287 CAGTCTGGCCCTGACGCGCCCATCGAAATTGTCGCGCGGATTAAATCTCGGCCCATCA 346
Db 608 CAGTCTGGCCCTGACGCGCCCATCGAAATTGTCGCGCGGATTAAATCTCGGCCCATCA 549

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds
(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacaccatcgaatgtgc.....acgacgttcaggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/FCI_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	10	US-09-765-555-16
2	3216.8	97.5	3300	10	US-09-765-555-18
3	3207.2	97.2	3300	10	US-09-765-555-14
4	3207.2	97.2	3300	10	US-09-765-555-15
5	3207.2	97.2	3300	10	US-09-765-555-17
6	2705.8	82.0	7373	18	US-10-257-384A-1
7	2705.8	82.0	8101	17	US-10-149-472-5

8	2701	81.8	6648	19	US-10-343-859-8	Sequence 8, Appli
9	2698	81.8	5191	19	US-10-343-859-9	Sequence 9, Appli
10	2690.4	81.5	7112	18	US-10-263-153-40	Sequence 40, Appl
11	2690.4	81.5	7259	18	US-10-263-153-30	Sequence 30, Appl
12	2690.4	81.5	7322	18	US-10-263-153-35	Sequence 35, Appl
13	2690.4	81.5	7352	18	US-10-263-153-25	Sequence 25, Appl
14	2690.4	81.5	7370	18	US-10-263-153-20	Sequence 20, Appl
15	2690.4	81.5	7370	18	US-10-263-153-61	Sequence 61, Appl
16	2690.4	81.5	7370	18	US-10-263-153-66	Sequence 66, Appl
17	2690.4	81.5	7370	18	US-10-263-153-71	Sequence 71, Appl
18	2690.4	81.5	7403	18	US-10-263-153-15	Sequence 15, Appl
19	2690.4	81.5	7442	18	US-10-263-153-10	Sequence 10, Appl
20	2690.4	81.5	7478	18	US-10-263-153-3	Sequence 3, Appli
21	2636.8	79.9	6806	10	US-09-908-943A-194	Sequence 194, App
22	2636.8	79.9	6806	20	US-10-801-487-194	Sequence 194, App
23	2636.8	79.9	6806	20	US-10-801-938-194	Sequence 194, App
24	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
25	2636.8	79.9	6806	20	US-10-801-486-194	Sequence 194, App
26	2636.8	79.9	6806	21	US-10-801-493-194	Sequence 194, App
27	2605.4	79.0	7553	18	US-10-263-153-7	Sequence 7, Appli
28	1609.4	48.8	5558	16	US-10-241-596-137	Sequence 137, App
29	1474	44.7	4700	17	US-10-313-963A-54	Sequence 54, Appl
30	1462.4	44.3	4921	21	US-10-495-491-2	Sequence 2, Appli
31	1462.4	44.3	4935	21	US-10-495-491-1	Sequence 1, Appli
32	1462.4	44.3	4945	21	US-10-495-491-3	Sequence 3, Appli
33	1462.4	44.3	4951	21	US-10-495-491-4	Sequence 4, Appli
34	1417.6	43.0	1511	10	US-09-987-763-30	Sequence 30, Appl
35	1417.6	43.0	5926	13	US-10-024-809-3	Sequence 3, Appli
36	1238.4	37.5	5024	17	US-10-313-963A-52	Sequence 52, Appl
37	1238.4	37.5	5448	21	US-10-671-995A-2	Sequence 2, Appli
38	1238.4	37.5	5454	21	US-10-671-995A-1	Sequence 1, Appli
39	1237.2	37.5	1922	10	US-09-995-988A-32	Sequence 32, Appl
40	1237.2	37.5	1922	17	US-10-420-034A-32	Sequence 32, Appl
41	1237.2	37.5	6270	21	US-10-855-897-2	Sequence 2, Appli
42	1204	36.5	4557	17	US-10-457-372-5	GENERAL INFORMA
43	1204	36.5	6664	18	US-10-270-176-1	Sequence 1, Appli
44	1204	36.5	6668	18	US-10-270-176-2	Sequence 2, Appli
45	1204	36.5	6742	18	US-10-270-176-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-765-555-16
; Sequence 16, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m3 and zinc finger
; OTHER INFORMATION: protein 2Fpm3
US-09-765-555-16

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACACCATCGAATGGTCAAAACCTTTCCGGGTATGGCATGTAGCGCCCGGAAGAGA 60
Db 1 CGGACACCATCGAATGGTCAAAACCTTTCCGGGTATGGCATGTAGCGCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATGCGG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATGCGG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCGGGTGGTGAACCGAGCCAGCCAGCTTTCTCGGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCGGGTGGTGAACCGAGCCAGCCAGCTTTCTCGGAAA 180
QY 181 CCGGGGAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
Db 181 CCGGGGAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
QY 241 AACAACTGCGGGGCAAAAGTGTGCTGATTTGGCGTGGCCACTTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGCGGGGCAAAAGTGTGCTGATTTGGCGTGGCCACTTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTTGTCGGGCGATTTAAATCTCGCGCGGATCAAATGGGTGCGACG 360
Db 301 ACGCGCGTTCGCAAAATTTGTCGGGCGATTTAAATCTCGCGCGGATCAAATGGGTGCGACG 360
QY 361 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCGCTGAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCGCTGAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCGCGGATCAAATGGGTGCGAC 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCGCGGATCAAATGGGTGCGAC 480
QY 481 TTGCTGTGGAAGTGCCTGCATTAATGTTCCGGGCTTATTTCTTGATGCTCTGACACAGA 540
Db 481 TTGCTGTGGAAGTGCCTGCATTAATGTTCCGGGCTTATTTCTTGATGCTCTGACACAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGGATCGGCGGTGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGGATCGGCGGTGAGCATC 600
QY 601 TGGTGCATTTGGTCCACGCAAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGCTCTCGG 660
Db 601 TGGTGCATTTGGTCCACGCAAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGCTCTCGG 660
QY 661 CGCGTCTGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db 661 CGCGTCTGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
QY 721 CGGAAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCGCAACGATCAGATGGCGCGCA 840
Db 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCGCAACGATCAGATGGCGCGCA 840
QY 841 TGGCGGCATTTACGAGTCCGGCTCGCGTGTGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCATTTACGAGTCCGGCTCGCGTGTGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCCATCAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAACCGAGGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAACCGAGGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCGTCTCACTGTGGAAGAAAACCACTTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCGTCTCACTGTGGAAGAAAACCACTTGGCGCCCAATA 1080
QY 1081 CGCAACCGGCTCTTCCCGCGCGTGGCCGATTTAATTAATGAGCTGGCAGCAGGTTT 1140

Db 1081 CGCAACCGGCTCTTCCCGCGCGTGGCCGATTTAATTAATGAGCTGGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAAGCAACCGCAATTAATGTAGTTAGCTCAGTCATTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAAGCAACCGCAATTAATGTAGTTAGCTCAGTCATTAG 1200
QY 1201 GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCCATCGGAAGCTGTGTATGGCTGTCAGGTGCTGAATCACTGCAATAATTCG 1320
Db 1261 TCAGCAGCCCATCGGAAGCTGTGTATGGCTGTCAGGTGCTGAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCCGTTCTCGATAAATGTTTTTTCGCGCGACATCAATAACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTCGATAAATGTTTTTTCGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATGAGCTGTTGAACAATTAATCATCGGCTCGTATAATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGAACAATTAATCATCGGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGGCTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGGCTGTTTTCACGA 1500
QY 1501 GCATCTTCCAAACGAGCAATAGATTATGAAGAACTGAAGAGGTAACTGGTAAATCTCG 1560
Db 1501 GCATCTTCCAAACGAGCAATAGATTATGAAGAACTGAAGAGGTAACTGGTAAATCTCG 1560
QY 1561 ATTAAACGGCGATTAAGGCTTATACGCTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGCGATTAAGGCTTATACGCTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGAAGAGAAATTCGAGAAAGAT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGAAGAGAAATTCGAGAAAGAT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCACCCGGCAAGAGCTTCCAGACAGCTGAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCACCCGGCAAGAGCTTCCAGACAGCTGAT 1800
QY 1801 CGTTTACCTGGGATGCCGTACGTTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CGTTTACCTGGGATGCCGTACGTTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTTAAACAGATCTGCTGCCAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTAAACAGATCTGCTGCCAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGAGGTAAGAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGAGGTAAGAGCGCGCTGATGTTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTTGGCGCTGATGCTGTGACGGGGTTTATCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTTGGCGCTGATGCTGTGACGGGGTTTATCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTTAAAGACGTGGCGTGAATACCGTGGCGGAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTTAAAGACGTGGCGTGAATACCGTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGATTAATAAACAATGAATGCAACACCGATGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGATTAATAAACAATGAATGCAACACCGATGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACCGCCCGTGG 2220

```
Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGCTGG 2220
Qy 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTATGGTGTAAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTATGGTGTAAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTCTTGGCGGTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTTGGCGGTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTCACTGATGAAGGCTCG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTCACTGATGAAGGCTCG 2400
Qy 2401 GAAGCGGTATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGAAAG 2460
Db 2401 GAAGCGGTATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGAAAG 2460
Qy 2461 TTGGGGAAGATCCACGTTATTCGCCGCCACCATGGAAAAAGCCACAAAAGGTGAATCATG 2520
Db 2461 TTGGGGAAGATCCACGTTATTCGCCGCCACCATGGAAAAAGCCACAAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTCTTGGTATGCCGTGCGTACTGCGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTCTTGGTATGCCGTGCGTACTGCGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGTCAGACTGTCGATGAAGCCCTGAAAGACGGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGACTGTCGATGAAGCCCTGAAAGACGGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACACACAAATAACATACACAACTCGGGATCGAGGGAAGGATTTTCAGAAATC 2700
Db 2641 AACACACACAAATAACATACACAACTCGGGATCGAGGGAAGGATTTTCAGAAATC 2700
Qy 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCGGGAGAGCCCTATCTGTT 2760
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCGGGAGAGCCCTATCTGTT 2760
Qy 2761 CCGGAATGTGTAAGTCTTTCAGCGATCTGTCGCCACCTGTTTTCAGCCAGCGGCTCC 2880
Db 2761 CCGGAATGTGTAAGTCTTTCAGCGATCTGTCGCCACCTGTTTTCAGCCAGCGGCTCC 2880
Qy 2881 ACGGTTGAANAACCGTATAAATGCCAGAGTGGCGCAATCTTTTACACAGCGGCTCC 2880
Db 2881 ACGGTTGAANAACCGTATAAATGCCAGAGTGGCGCAATCTTTTACACAGCGGCTCC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGATGTGGC 2940
Qy 2941 AAGTCTTTTCAGCCAGAGCTCCAGCCTGGTGGCCACCAACGTAATCTTTCAGCCAGGAG 3000
Db 2941 AAGTCTTTTCAGCCAGAGCTCCAGCCTGGTGGCCACCAACGTAATCTTTCAGCCAGGAG 3000
Qy 3001 CCCTATGCTTGTGCGGAATGTGGTAAGTCTTTCAGCCAGAGAGCTCCCTGTTGGCCAC 3060
Db 3001 CCCTATGCTTGTGCGGAATGTGGTAAGTCTTTCAGCCAGAGAGCTCCCTGTTGGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAAACCGTATAATGCCAGAGTGGCGAAATCTTTTAGT 3120
Db 3061 CAGCGTACCCACACGGGTGAAAACCGTATAATGCCAGAGTGGCGAAATCTTTTAGT 3120
Qy 3121 GACTGCCGCGACTTGTGCGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 GACTGCCGCGACTTGTGCGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Qy 3181 CCAGAAATGTGGCAAGTCTTTTCTCCAAATCCAGCCATCTGCTCCGGACCAACGTAATCAT 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTTCTCCAAATCCAGCCATCTGCTCCGGACCAACGTAATCAT 3240
Qy 3241 ACCGGTAAAAAATAGTGGCCAGCGCGAGTACCGGTACGAGTTCGGGACTACCGT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGCGCGAGTACCGGTACGAGTTCGGGACTACCGT 3300
```

RESULT 2

```
US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Parial sequence of pMal-Ap3 and zinc finger
; OTHER INFORMATION: protein ZFPap3
US-09-765-555-18
```

Query Match 97.5%; Score 3216.8; DB 10; Length 3300;

Best Local Similarity 98.4%; Pred. No. 0;

Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60

Db 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60

Qy 61 GTCAATTTCAGGGTGTGATGTGAACCAAGTAAAGTATACGATGTCGACAGTATGCGG 120

Db 61 GTCAATTTCAGGGTGTGATGTGAACCAAGTAAAGTATACGATGTCGACAGTATGCGG 120

Qy 121 GTGTCCTTATCAGACCCGTTTCCCGCGTGTGAACCCAGCGCAGCCACGTTTCTCGGAAAA 180

Db 121 GTGTCCTTATCAGACCCGTTTCCCGCGTGTGAACCCAGCGCAGCCACGTTTCTCGGAAAA 180

Qy 181 CCGGGAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTACATTCCTCAACCGCGTGGCAC 240

Db 181 CCGGGAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTACATTCCTCAACCGCGTGGCAC 240

Qy 241 AACAACTCGCGGGCAACAGTCTGTTGCTGATGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300

Db 241 AACAACTCGCGGGCAACAGTCTGTTGCTGATGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300

Qy 301 ACGGCGCTGCGCAATTCGCGCGGATTAATCTCGCGCGCATCACTCGGTCGCGGCG 360

Db 301 ACGGCGCTGCGCAATTCGCGCGGATTAATCTCGCGCGCATCACTCGGTCGCGGCG 360

Qy 361 TGGTGGTTCGATGTTAGAACGAGCGCGCTCGAAGCCCTGAAAGCGCGGTCGCAATC 420

Db 361 TGGTGGTTCGATGTTAGAACGAGCGCGCTCGAAGCCCTGAAAGCGCGGTCGCAATC 420

Qy 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCGCTGGATGATCAAGGATGCCA 480

Db 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCGCTGGATGATCAAGGATGCCA 480

Qy 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTATTTCTTCATGTTCTGACCCAGA 540

Db 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTATTTCTTCATGTTCTGACCCAGA 540

Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600

Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600

Qy 601 TGGTCCGATTTGGGTCCACAGCAAAATCGCGCTGTTTAGCGGGGCCCATTAAGTTCTGTCGG 660

Db 601 TGGTCCGATTTGGGTCCACAGCAAAATCGCGCTGTTTAGCGGGGCCCATTAAGTTCTGTCGG 660

Db 601 |||||TGTCGCAATTGGTTCACACGAAATCGCGCTGTGTAGCGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CCGCTCTGGTCTGGCTGGCTGGCTGCAATAATATCTCTCTGCAATCAAAATTCAGCGGATAG 720
Db 661 CCGCTCTGGCTGGCTGGCTGGCTGCAATAATATCTCTCTGCAATCAAAATTCAGCGGATAG 720
QY 721 CCGAAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CCGAAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCTGTTCCCATCTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCTGTTCCCATCTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGAA 840
QY 841 TCGCGGCCAATACCGAGTCCGGCTGGCGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCCAATACCGAGTCCGGCTGGCGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGTCGCCGCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGTCGCCGCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCAGCTGSCACGACAGTTT 1140
Db 1081 CGAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCAGCTGSCACGACAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTCGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTCGCG 1260
QY 1261 TCAGGACGCAATCGAAGCTGTGTATGCTGCTGCAAGTGTAAATCACTGCAATAATTCG 1320
Db 1261 TCAGGACGCAATCGAAGCTGTGTATGCTGCTGCAAGTGTAAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACCGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACCGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTGTGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTGTGA 1440
QY 1441 ATTGTAGCGGATAAACAATTTACACAGGAAACAGCGCTCGGTTAGGTGTTTACGA 1500
Db 1441 ATTGTAGCGGATAAACAATTTACACAGGAAACAGCGCTCGGTTAGGTGTTTACGA 1500
QY 1501 GCACTTCAACAAAGGACCATAGATATGAAACCTGAAGAAGTAACTGGTAAATCTGG 1560
Db 1501 GCACTTCAACAAAGGACCATAGATATGAAACCTGAAGAAGTAACTGGTAAATCTGG 1560
QY 1561 ATTAACGGGATTAAGGCTATACGTTCTGCTGAAGTCGGTGAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGGATTAAGGCTATACGTTCTGCTGAAGTCGGTGAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATTAACCTGGAAGAAATTCGCAAGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATTAACCTGGAAGAAATTCGCAAGTT 1680
QY 1681 GCGGCAATCGGAGTGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
|||||

Db 1681 GCGGCAATCGGAGTGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGACAACTGAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGACAACTGAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTAACAAGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTAACAAGCAAGCTGATGCTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATACAAGATCTGCTGCCGAACCCGCCAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATACAAGATCTGCTGCCGAACCCGCCAAACCTCGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGATGCTGATTTAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGATGCTGATTTAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAAGCTGCTTTTAAAGCGAAACAGCGATGACCATCAACGCCCGGTGG 2220
Db 2161 TCCATCGCAAGCTGCTTTTAAAGCGAAACAGCGATGACCATCAACGCCCGGTGG 2220
QY 2221 GCATGTTCCAACTGCACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC 2280
Db 2221 GCATGTTCCAACTGCACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC 2280
QY 2281 AAGGTTCAACCATCCAAACCGTTGTTGGGTGCTGAGCGCAGGTATTAAACCGCCAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTGTTGGGTGCTGAGCGCAGGTATTAAACCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTAATAAGACAAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTCGCCCAACCATGGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTCGCCCAACCATGGAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCGAGGCTGCTGAGCTGTCGATGAAGCCTGAAAGACCGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCGAGGCTGCTGAGCTGTCGATGAAGCCTGAAAGACCGCGCAGACTAAATTCGAGTCTG 2640
QY 2641 AACAAACAAACAAATAAACAACCACTCGGATTCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAAACAAATAAACAACCACTCGGATTCGAGGGAAGGATTTCAAGATTC 2700
QY 2701 GGATCTCTTCTGTCGCGCAGCGCCCTCGAGCCCGGGGAAGCCCTATCTTGT 2760
Db 2701 GGATCTCTTCTGTCGCGCAGCGCCCTCGAGCCCGGGGAAGCCCTATCTTGT 2760
QY 2761 CCGGAATGTGGTAAGTCTTCAAGCATCTCGGCCACTGTTTCGCCACAGGTTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTCAAGCATCTTCCCTGGTGGTCCACAGGTTACCCAC 2820

QY 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTTAGCACACAGCGGTCC 2880
Db |||||
QY 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
Db |||||
QY 2881 CTGGTGGCCATCAACGCACTCATCTGGCGAGAGCCATACAAATGTCCAGAAATGTGC 2940
Db |||||
QY 2881 CTGGTGGCCATCAACGCACTCATCTGGCGAGAGCCATACAAATGTCCAGAAATGTGC 2940
Db |||||
QY 2941 AAGTCTTTTCAGCAGAGCTCCAGCCTGGTGGCGCACCAACGTACTCACACCGGGGAGAAG 3000
Db |||||
QY 2941 AAGTCTTTTCAGCAGAGCTCCAGCAGCAACCTGGTGGCGCACCAACGTACTCACACCGGGGAGAAG 3000
Db |||||
QY 3001 CCCTATGCTTGTCCGGAATGTGAAGTCTTTCAGCCAGAGAGCAGTCCCTGGTGGCGCAC 3060
Db |||||
QY 3001 CCCTATGCTTGTCCGGAATGTGAAGTCTTTCAGCAGCAGAGTGGTCTCTTGGTTAGACAC 3060
Db |||||
QY 3061 CAGCGTACCCACACCGGGTGAAAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTTAGT 3120
Db |||||
QY 3061 CAGCGTACCCACACCGGGTGAAAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db |||||
QY 3121 GACTGCCGCGACTTGTCTCGCCATCAACGCACTCATCTGGCGAGAGCCATACAAATGT 3180
Db |||||
QY 3121 CAGCGCGCCACTTGGAGCGCATCAACGCACTCATCTGGCGAGAGCCATACAAATGT 3180
Db |||||
QY 3181 CCAGAAATGTGCAAGTCTTTTCTCCAAATCCAGCCATCTGCTCCGGCACCAAGTACTCAC 3240
Db |||||
QY 3181 CCAGAAATGTGCAAGTCTTTTCTCAACTTCAGGCAACTTGGTCCGTCAACCAAGTACTCAC 3240
Db |||||
QY 3241 ACCGGTAAAAAAGTGTGGCCAGCGCGCCAGTACCGGTACGAGTTCGGGACTACGCT 3300
Db |||||
QY 3241 ACCGGTAAAAAAGTGTGGCCAGCGCGCCAGTACCGGTACGAGTTCGGGACTACGCT 3300
Db |||||

RESULT 3

US-09-765-555-14
; Sequence 14, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Partial sequence of pMal-m1 and zinc finger
; OTHER INFORMATION: protein 2FPm1

US-09-765-555-14

Query Match 97.2%; Score 3207.2; DB 10; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 1 CCGACACCATCAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60
Db |||||
QY 1 CCGACACCATCAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60
Db |||||
QY 61 GTCAATTCAAGGTGGTGAATGTGAACAGTAACTGTTATACCATGTTCGAGAGTATGCG 120
Db |||||
QY 61 GTCAATTCAAGGTGGTGAATGTGAACAGTAACTGTTATACCATGTTCGAGAGTATGCG 120
Db |||||
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACAGGCCACGCTTCTCGCAAAA 180
Db |||||

Db |||||
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACAGGCCAGCCACGCTTCTCGCAAAA 180
QY |||||
Db 181 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
QY |||||
Db 181 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
QY |||||
QY 241 AACAACTGGCGGGGAAAAACAGTCTGTGATGTGGGTTGGCACTCTCAAGTCTGGCCCTGC 300
Db |||||
QY 241 AACAACTGGCGGGGAAAAACAGTCTGTGATGTGGGTTGGCACTCTCAAGTCTGGCCCTGC 300
Db |||||
QY 301 ACGCGCGTCCCAAAATGTCTCGCGCGATTAATCTCGCGCGATCAATCTGGGTGGCGCACG 360
Db |||||
QY 301 ACGCGCGTCCCAAAATGTCTCGCGCGATTAATCTCGCGCGATCAATCTGGGTGGCGCACG 360
Db |||||
QY 361 TGGTGGTGTGATGATAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db |||||
QY 361 TGGTGGTGTGATGATAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db |||||
QY 421 TTCTCGCGCAACCGCGTCAAGTGGGCTGATCAATACTATCCGCTGGATGACAGGATGCCA 480
Db |||||
QY 421 TTCTCGCGCAACCGCGTCAAGTGGGCTGATCAATACTATCCGCTGGATGACAGGATGCCA 480
Db |||||
QY 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCGGCGGTATTTCTTGTATCTCTGACCCAGA 540
Db |||||
QY 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCGGCGGTATTTCTTGTATCTCTGACCCAGA 540
Db |||||
QY 541 CACCCATCAACAGATATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db |||||
QY 541 CACCCATCAACAGATATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db |||||
QY 601 TGGTGGCATTTGGGTGACAGCAAAATTCGCGTGTGTAGCGGGCCCATTAAGATTTCTGTCGG 660
Db |||||
QY 601 TGGTGGCATTTGGGTGACAGCAAAATTCGCGTGTGTAGCGGGCCCATTAAGATTTCTGTCGG 660
Db |||||
QY 661 CGCGTCTCGGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db |||||
QY 661 CGCGTCTCGGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db |||||
QY 721 CCGAAACGGGAAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db |||||
QY 721 CCGAAACGGGAAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db |||||
QY 781 ATGAGGGCATCGTTCCTCCACTCGCATGTGTTGTGCAACAGATCAGATGCGCTGGCGCAA 840
Db |||||
QY 781 ATGAGGGCATCGTTCCTCCACTCGCATGTGTTGTGCAACAGATCAGATGCGCTGGCGCAA 840
Db |||||
QY 841 TCGCGGCCATTTACCGAGTCCCGGCTGGCGGTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db |||||
QY 841 TCGCGGCCATTTACCGAGTCCCGGCTGGCGGTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db |||||
QY 901 ACGATACCGAAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
Db |||||
QY 901 ACGATACCGAAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
Db |||||
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTCTCTCAAGTGGCGGCGGTGA 1020
Db |||||
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTCTCTCAAGTGGCGGCGGTGA 1020
Db |||||
QY 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAAAACCCCTTGGCGGCCAATA 1080
Db |||||
QY 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAAAACCCCTTGGCGGCCAATA 1080
Db |||||
QY 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGCGACGACAGGTTT 1140
Db |||||
QY 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGCGACGACAGGTTT 1140
Db |||||
QY 1141 CCGCACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTGTAGCTCACTCATTAG 1200
Db |||||
QY 1141 CCGCACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTGTAGCTCACTCATTAG 1200
Db |||||
QY 1201 GCACAATTTCTCATGTTTTCAGAGCTTATCATGCACTGCAAGTGGCAACCAATGCTTCTGGCG 1260
Db |||||

Db 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAGGTGCTGTAATCACTGCATCAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAGGTGCTGTAATCACTGCATCAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCGCTTCTGGATAATGTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCGCTTCTGGATAATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGACCTGTGACAAATTAATCATCGGCTGTAATGTTGCGA 1440
Db 1381 CTGGCAATATTTCTGAAATGACCTGTGACAAATTAATCATCGGCTGTAATGTTGCGA 1440
QY 1441 ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAAG 1500
Db 1441 ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAAG 1500
QY 1501 GCACCTCACCAACAGGACCATAGATTATGAAACTGGAAGAGTAACTCGTGAATCTGG 1560
Db 1501 GCACCTCACCAACAGGACCATAGATTATGAAACTGGAAGAGTAACTCGTGAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTTAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTTAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATCCCAAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATCCCAAGGTT 1680
QY 1681 GCGCACTGCGGATGCGCTTGAATTAACCCGGACAAAGCGTTCAGGACAAAGCTGAT 1740
Db 1681 GCGCACTGCGGATGCGCTTGAATTAACCCGGACAAAGCGTTCAGGACAAAGCTGAT 1740
QY 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGATGCGCTTGAATTAACCCGGACAAAGCGTTCAGGACAAAGCTGAT 1860
Db 1801 CCGTTTACCTGGATGCGCTTGAATTAACCCGGACAAAGCGTTCAGGACAAAGCTGAT 1860
QY 1861 GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCGAAGAGTAAAGCGCGCTGATTTCAAC 1920
Db 1861 GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCGAAGAGTAAAGCGCGCTGATTTCAAC 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGTCGAGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGTCGAGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGCGTGGATACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGCGTGGATACGCTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGATGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCGCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCGCGTGG 2220
QY 2221 GCATGGTCCACATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGGCTGCTGAGCGCAGGTATTAAACGGCGCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGTTGGGCTGCTGAGCGCAGGTATTAAACGGCGCAGT 2340

RESULT 4

US-09-765-555-15

; Sequence 15, Application US/09765555

; Publication No. US20030037355A1

; GENERAL INFORMATION:

; APPLICANT: The Scripps Research Institute

; TITLE OF INVENTION: Methods and compositions to modulate

; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger
; OTHER INFORMATION: protein ZFPm2
US-09-765-555-15

Query Match 97.2%; Score 3207.2; DB 10; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY	1	CGGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA	60
Db	1	CGGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA	60
QY	61	GTCAATTACAGGTGGTGAATGTGA	AAACAGTAACTTATACGATGTGCGAGAGTATGCG	120
Db	61	GTCAATTACAGGTGGTGAATGTGA	AAACAGTAACTTATACGATGTGCGAGAGTATGCG	120
QY	121	GTGTCTCTTATCAGACCGTTTCC	CGGTGTAACACGAGCCACGCTTTCTCGGAAAA	180
Db	121	GTGTCTCTTATCAGACCGTTTCC	CGGTGTAACACGAGCCACGCTTTCTCGGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGGATG	GGGAGCTGAATTACATTTCCCAACCGGTGGCAC	240
Db	181	CGCGGAAAAAGTGAAGCGGATG	GGGAGCTGAATTACATTTCCCAACCGGTGGCAC	240
QY	241	AACAACTGGCGGCAAAACAGT	CTGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC	300
Db	241	AACAACTGGCGGCAAAACAGT	CTGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC	300
QY	301	ACGCGCGTTCGCAATTTGTCGG	CGCATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCGTTCGCAATTTGTCGG	CGCATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
QY	361	TGGTGTGTGATGTAGAACGA	AGGGCGTTCGAAGCCTGTAAAGCGGCGTGCACAATC	420
Db	361	TGGTGTGTGATGTAGAACGA	AGGGCGTTCGAAGCCTGTAAAGCGGCGTGCACAATC	420
QY	421	TTCTCGCGCAACCGTTCAGT	GGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
Db	421	TTCTCGCGCAACCGTTCAGT	GGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCTGCA	CTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA	540
Db	481	TTGCTGTGGAAGCTGCTGCA	CTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTATTT	CTCCATGAAGACGTTAGCGGCTGGAGGCATC	600
Db	541	CACCCATCAACAGTATTATTT	CTCCATGAAGACGTTAGCGGCTGGAGGCATC	600
QY	601	TGGTTCGATGGTTCACAG	CAAACTGCTGTAGCGGGCCCATTAAGTCTGTCTCGG	660
Db	601	TGGTTCGATGGTTCACAG	CAAACTGCTGTAGCGGGCCCATTAAGTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGG	CAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGCGTCTGGCTGG	CAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGCATGTG	CGATGTCGCGTTTTCAAACCAATGCAAAATGCTGA	780
Db				

Db	721	CGGAACGGGAAGCGCATGTG	AGTGCATGTCCGGTTTTCAAACCAATGCAAAATGCTGA	780
QY	781	ATGAGGCGCATCGTTCCCA	CTCGATGTGTTGGCAACGATCAGATGCGCTGGCGCGCAA	840
Db	781	ATGAGGCGCATCGTTCCCA	CTCGATGTGTTGGCAACGATCAGATGCGCTGGCGCGCAA	840
QY	841	TGCGGCGCATTAACGAGT	CCGGCTGCGCGTGGTGGGATATCTCGGTAGTGGGATAG	900
Db	841	TGCGGCGCATTAACGAGT	CCGGCTGCGCGTGGTGGGATATCTCGGTAGTGGGATAG	900
QY	901	ACGATACCGAAGACAGCT	CATGTTATATCCCGCGTTAAACCAACATCAAAACGAGATTTTC	960
Db	901	ACGATACCGAAGACAGCT	CATGTTATATCCCGCGTTAAACCAACATCAAAACGAGATTTTC	960
QY	961	GCCTGCTGGGGCAAAAC	CAGCGCTTGTGCTCAACTCTCTCAGGGCCAGGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAAC	CAGCGCTTGTGCTCAACTCTCTCAGGGCCAGGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGCC	CGTCTCACTGTGTAAGAAAAACCAACCTGGCGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCC	CGTCTCACTGTGTAAGAAAAACCAACCTGGCGCCCAATA	1080
QY	1081	CGAAACCGCTCTCCCG	CGGCTTGCAGTTCATTAATGCAGCTGCACGACAGGTTT	1140
Db	1081	CGAAACCGCTCTCCCG	CGGCTTGCAGTTCATTAATGCAGCTGCACGACAGGTTT	1140
QY	1141	CCGACTGGAAGGGG	CAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGACTGGAAGGGG	CAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAATTTCTATGTTG	ACAGCTTATCATGCTGCAAGGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACAATTTCTATGTTG	ACAGCTTATCATGCTGCAAGGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGCAGCCATCGGA	AGCTGTGATGCTGAGTTCGAGTTCGTAATCACTGCATTAATTCG	1320
Db	1261	TCAGCAGCCATCGGA	AGCTGTGATGCTGAGTTCGAGTTCGTAATCACTGCATTAATTCG	1320
QY	1321	TGTGCTCAAGCG	CACTCCCGTCTCGATTAATGTTTTTGGCGCGACATCAATAACGGTT	1380
Db	1321	TGTGCTCAAGCG	CACTCCCGTCTCGATTAATGTTTTTGGCGCGACATCAATAACGGTT	1380
QY	1381	CTGGCAAAATTTCTG	AAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA	1440
Db	1381	CTGGCAAAATTTCTG	AAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA	1440
QY	1441	ATTGTGAGCGGAT	AACAATTTTACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA	1500
Db	1441	ATTGTGAGCGGAT	AACAATTTTACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA	1500
QY	1501	GCACCTTCAACCA	AGGACCATAGATTATGAAACTGGAAGAGGTAAACTTGGTAATCTGG	1560
Db	1501	GCACCTTCAACCA	AGGACCATAGATTATGAAACTGGAAGAGGTAAACTTGGTAATCTGG	1560
QY	1561	ATTAAACGCGAT	ATAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGCGAT	ATAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGT	CACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT	1680
Db	1621	ACCGGAATTAAGT	CACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT	1680
QY	1681	GGCGCAACTGCG	CGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	GGCGCAACTGCG	CGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCG	CTGTTGGCTGAATCAACCCGGGAAGCGTTCCAGGAACAGCTGAT	1800
Db	1741	GCTCAATCTGCG	CTGTTGGCTGAATCAACCCGGGAAGCGTTCCAGGAACAGCTGAT	1800
QY	1801	CGGTTTACCTGG	ATGCGTACGTTTACACCGCAAGCTGATGCTTACCCGATCCCTGTT	1860
Db	1801	CGGTTTACCTGG	ATGCGTACGTTTACACCGCAAGCTGATGCTTACCCGATCCCTGTT	1860

1861 GAAGCGTTATCGTGATTTTATAACAAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
1861 GAAGCGTTATCGTGATTTTATAACAAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTGAAGCGCGCTGATTTCAAC 1980
1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTGAAGCGCGCTGATTTCAAC 1980
1981 CTGCAAGAACCTGACTTACCTCGCGCTGATTTGCTGCTGACGCGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCTGACTTACCTCGCGCTGATTTGCTGCTGACGCGGGTTATGCGTTCAAG 2040
2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACCGCTGGCGCGAAACGG 2100
2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACCGCTGGCGCGAAACGG 2100
2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAGCAAAACACATGAATGCGAGACACCGATTAC 2160
2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAGCAAAACACATGAATGCGAGACACCGATTAC 2160
2161 TCCATCGCAGAACCTGCTTTTAAAGCGGCGTGGATAACCGCTGGCGCGAAACGG 2220
2161 TCCATCGCAGAACCTGCTTTTAAAGCGGCGTGGATAACCGCTGGCGCGAAACGG 2220
2221 GCATGGTCCAAATCGACACCAAGCAAGTGAATTAAGTGAACGGTACTGCCGACCTTC 2280
2221 GCATGGTCCAAATCGACACCAAGCAAGTGAATTAAGTGAACGGTACTGCCGACCTTC 2280
2281 AAGGTCACCAATCCAAACCGTTCGTCGCGCTGAGCGAGGATTAAGCGCGCCAGT 2340
2281 AAGGTCACCAATCCAAACCGTTCGTCGCGCTGAGCGAGGATTAAGCGCGCCAGT 2340
2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
2401 GAAGCGTTTAAAGACAAACCGCTGGGTCGCTGAGCTGAAAGTCTTACGAGGAAGAG 2460
2401 GAAGCGTTTAAAGACAAACCGCTGGGTCGCTGAGCTGAAAGTCTTACGAGGAAGAG 2460
2461 TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAACCGCCAGAAAGGTAATCATG 2520
2461 TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAACCGCCAGAAAGGTAATCATG 2520
2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC 2580
2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC 2580
2581 GCCAGCGTCTGACGCTGATGAGCCCTGGAAGCGCGCAGACTAATTCGAGCTCG 2640
2581 GCCAGCGTCTGACGCTGATGAGCCCTGGAAGCGCGCAGACTAATTCGAGCTCG 2640
2641 AACAAACAAACAAATAACAAACAACTCGGGATCGAGGGAAGGATTTAGAAATTC 2700
2641 AACAAACAAACAAATAACAAACAACTCGGGATCGAGGGAAGGATTTAGAAATTC 2700
2701 GGATCTCTTCTCTGTCGCGCGCTTCGAGCGCGGGGAGAGCCCTATGCTTGT 2760
2701 GGATCTCTTCTCTGTCGCGCGCTTCGAGCGCGGGGAGAGCCCTATGCTTGT 2760
2761 CCGGAATGTGTAGTCTCTTCAGCGATCTGGCCACCTGCTGCGCACCGCTACCCAC 2820
2761 CCGGAATGTGTAGTCTCTTCAGAGCTCTCACCTGCTGCGCGCCACCGCTACCCAC 2820
2821 ACGGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCAGCAGCGCTCC 2880
2821 ACGGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCAGCAGCGCTCC 2880
2881 CTGTCGCGCATCAACGCTCTACTCTGCGGAGGACCATACAAATGTCAGAAATGTGGC 2940
2881 CTGTCGCGCATCAACGCTCTACTCTGCGGAGGACCATACAAATGTCAGAAATGTGGC 2940

2941 AAGTCTTTACGCCAGAGCTCAGCGCTGGTGGCCCAACAAAGTACTCTACACCGGGGAGAAG 3000
2941 AAGTCTTTCTCGGTCTGCAATCTCGTCGGCACCACAAAGTACTCTACACCGGGGAGAAG 3000
3001 CCCTATGCTTCTCGGGAATGTGTAAGTCTCTCAGCCAGAGCAGCTCCCTGGTGGCCAC 3060
3001 CCCTATGCTTCTCGGGAATGTGTAAGTCTCTCAGCCAGAGCAGTAACTGGTGGCCAC 3060
3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAAATGCCAGAGTGGCGAAATCTTTTGT 3120
3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAAATGCCAGAGTGGCGAAATCTTTTGT 3120
3121 GACTGCCGCGACCTTCTCGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGT 3180
3121 CAGCCCGGCGACCTTCTCGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGT 3180
3181 CCAGAAATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTCGTCGGCTGCAATCTCGTCGGCACCAACGTAATGT 3240
3181 CCAGAAATGTGGCAAGTCTTTCTCGCTGCAATCTCGTCGGCACCAACGTAATGT 3240
3241 ACCGGTAAAAAATACTAGTGGCGCGCGCGCGAGTACCGGTACGAGTTCGGGACTACGCT 3300
3241 ACCGGTAAAAAATACTAGTGGCGCGCGCGCGAGTACCGGTACGAGTTCGGGACTACGCT 3300

RESULT 5
US-09-765-555-17
; Sequence 17, Application US/09765555
; Publication No. US2003003735A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein ZPFm4
US-09-765-555-17

Query Match 97.2%; Score 3207.2; DB 10; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

1 CCGACACCATCGAATGGTGCATAAAACCTTTTCGGGGTATGCGATGATAGCGCCGGAAGAGA 60
1 CCGACACCATCGAATGGTGCATAAAACCTTTTCGGGGTATGCGATGATAGCGCCGGAAGAGA 60
61 GTCAATTTCAGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTCGAGAGTATGCCG 120
61 GTCAATTTCAGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTCGAGAGTATGCCG 120
121 GTGCTCTTTATCAGACCGCTTTCCCGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTCGAGAGTATGCCG 180
121 GTGCTCTTTATCAGACCGCTTTCCCGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTCGAGAGTATGCCG 180
181 CCGCGGAAAAAGTGAAGCGCGGATGGCGAGTGAATTTACATTTCCCAACCGCGTGGCAC 240
181 CCGCGGAAAAAGTGAAGCGCGGATGGCGAGTGAATTTACATTTCCCAACCGCGTGGCAC 240
241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGGCTTGCACCTCCAGTCTGGCCTGCG 300

Db 241 AACAACTGGCGGCAAA CAGTGGTTCGTAATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 AGCGCGCTGCGAAATGTCGCGCGAATAAATCTCGCGCCGATCAACTGGGTGCCAGC 360
Db 301 AGCGCGCTGCGAAATGTCGCGCGAATAAATCTCGCGCCGATCAACTGGGTGCCAGC 360
Qy 361 TGGTGGTGCAGTGTAGAAACGAAGCGGTCGAGCCCTGTAAAGCGGCGTGCACAATC 420
Db 361 TGGTGGTGCAGTGTAGAAACGAAGCGGTCGAGCCCTGTAAAGCGGCGTGCACAATC 420
Qy 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCCTGCATAAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCATAAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGGCATTTGGTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTCGCTGGCTGGCTGCGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGGCTGGCTGCGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTCGCAACGATCGAGTGGCGTGGCGCA 840
Db 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTCGCAACGATCGAGTGGCGTGGCGCA 840
Qy 841 TGCGCGCATTAACGAGTCCGGCTGCGCGTTGGTGGGATATCTCGTAGTGGGATAG 900
Db 841 TGCGCGCATTAACGAGTCCGGCTGCGCGTTGGTGGGATATCTCGTAGTGGGATAG 900
Qy 901 AGGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
Db 901 AGGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
Qy 1081 GCGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGGTTT 1140
Db 1081 GCGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGAAATTAATGTGAGTTAGCTACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGAAATTAATGTGAGTTAGCTACTCATTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATGACTGCAAGGTCGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATGACTGCAAGGTCGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCAATATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCAATATTCG 1320
Qy 1321 TGTCCGCTCAAGGGCACTCCCGTTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTCCGCTCAAGGGCACTCCCGTTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380

Qy 1381 CTGGCAAAATATTTCTGAATAGCTGTTGCAAAATTAATCATCGGCTCGTATATATGTTGTGA 1440
Db 1381 CTGGCAAAATATTTCTGAATAGCTGTTGCAAAATTAATCATCGGCTCGTATATATGTTGTGA 1440
Qy 1441 ATTGTAGCGGATACCAATTTTCACAGGAACACGAGTCCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATACCAATTTTCACAGGAACACGAGTCCGTTTAGGTGTTTTCACGA 1500
Qy 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTTATCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTTATCTGG 1560
Qy 1561 ATTTAAACGCGATAAAGGCTATAACGCTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAACGCGATAAAGGCTATAACGCTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCAGGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGAACAAGCGTTCCAGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGAACAAGCGTTCCAGACAAGCTGTAT 1800
Qy 1801 CCGTTTACTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACGCTACTTCACTGGCGCTGATTTGCTGCTGAGCGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCACTGGCGCTGATTTGCTGCTGAGCGGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGATTAACAGACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGATTAACAGACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAACCAACACATGAATGCAAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAACCAACACATGAATGCAAGACACCGATTAC 2160
Qy 2161 TCCATCGAGAAAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGAGAAAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACACCATCAAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAACTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAACTCTTACGAGGAAG 2460

QY 541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCGCATAG 720
Db 661 CGCGTCTCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCGCATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTCCCACTGGATGTCTGGTTGCCAACGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCGTCCCACTGGATGTCTGGTTGCCAACGATCAGATGGCGCTGGCGCGAA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTATATCCCGCCGTTAACCHACCATCAAAACGAGTTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATCCCGCCGTTAACCHACCATCAAAACGAGTTTTC 960
QY 961 GCCTCTGGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTCTCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTCTCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA 1500
QY 1501 GCATCTCACCAACAGGACCATAGATTATGAAAACCTGGAAGGTTAAATCTGGTATCTGG 1560
Db 1501 GCATCTCACCAACAGGACCATAGATTATGAAAACCTGGAAGGTTAAATCTGGTATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTTAAACGCTCTCGCTGAAAGTCCGTTAAGAAAATTCAGAAAAG 1620
Db 1561 ATTTAAGCGGATAAAGGCTTAAACGCTCTCGCTGAAAGTCCGTTAAGAAAATTCAGAAAAG 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGATAACTGGAAGAGAAATCCCACAGGTT 1680
QY 1681 CGCGCAACTGCGGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTTAC 1740
Db 1681 CGCGCAACTGCGGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCCGACAAGCGTTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCCGACAAGCGTTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTACAACGGCAAGCTGATTGCTTACCCTGCTCTGT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTACAACGGCAAGCTGATTGCTTACCCTGCTCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATATAAGAACTGAAAGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATATAAGAACTGAAAGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGATTAACAGACCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGATTAACAGACCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTTGAACCTGATTAATAAACAACAACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTTGAACCTGATTAATAAACAACAACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACACGCAAAAGTGAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACACGCAAAAGTGAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCATCAAAACCGTTGTTGGCTGCTGAGCGAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCATCAAAACCGTTGTTGGCTGCTGAGCGAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATGGAAGAAACGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATGGAAGAAACGCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGCTGACATGTCGTAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGACATGTCGTAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
Db 2641 AACCAACAACAATAAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
QY 2701 GGATCCTCTTCTGCTGGCCCGCGGCTCGAGGCGCG 2741

Db 2701 GGATCCTCAGTCCCGGCAAGCCGCGCGCTGGTGGGAGG 2741

RESULT 7
US-10-149-472-5
; Sequence 5 Application US/10149472
; Publication No. US20040029204A1
; GENERAL INFORMATION:
; APPLICANT: DOUBIN-GRAMATICA, FRANCOISE
; APPLICANT: GOUBIN-GRAMATICA, FRANCOISE
; APPLICANT: PREVOST, GREGOIRE
; TITLE OF INVENTION: METHOD FOR OBTAINING HUMAN CDC25 PHOSPHATASES AND METHOD
; TITLE OF INVENTION: FOR IDENTIFYING HUMAN CDC25 PHOSPHATASE MODULATORS
; FILE REFERENCE: bml-427.065
; CURRENT APPLICATION NUMBER: US/10/149,472
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: PCT/FR00/03496
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: FR 99/15722
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: FR 00/06883
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: FR 00/12008
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatencIn ver. 2.1
; SEQ ID NO 5
; LENGTH: 8101
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: JM109 / pMAL - Hs Cdc25C strain
US-10-149-472-5

Query Match 82.0%; Score 2705.8; DB 17; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGTATGGCATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGTATGGCATGATAGCGCCGGAAGAGA 60

QY 61 GTCGAATTCAGGGTGGTGAATGGAACGAGTAACTTATACGATTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCGAATTCAGGGTGGTGAATGGAACGAGTAACTTATACGATTATACGATGTCGAGAGTATGCCG 120

QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCGAGCCAGCCACGTTTCTGCGAATA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCGAGCCAGCCACGTTTCTGCGAATA 180

QY 181 CGCGGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240

QY 241 AACAACTGGCGGGCAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGCCCTGCG 300
Db 241 AACAACTGGCGGGCAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGCCCTGCG 300

QY 301 ACGCGCGTCCGCAATTTGTGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTCCGCAATTTGTGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360

QY 361 TGGTGGTTCGATGGTGAAGCGCGTTCGAAAGCGCGCTGTAAGCGCGCGTTCGCAATC 420
Db 361 TGGTGGTTCGATGGTGAAGCGCGTTCGAAAGCGCGCTGTAAGCGCGCGTTCGCAATC 420

QY 421 TTCTCGCGCAACCGGTTCAGTGGGCTGATCATTAATCTACGCTGGATGACCGAGTCCCA 480
Db 421 TTCTCGCGCAACCGGTTCAGTGGGCTGATCATTAATCTACGCTGGATGACCGAGTCCCA 480

QY 481 TTGCTGGGAAGCTGCCTGCACCTAAAGTTCCGCGTTATTTCTTGATGTTCTTGACGAGA 540
Db 481 TTGCTGGGAAGCTGCCTGCACCTAAAGTTCCGCGTTATTTCTTGATGTTCTTGACGAGA 540

Db 481 TTGCTGGGAAGCTGCCTGCACCTAAAGTTCCGCGTTATTTCTTGATGTTCTTGACGAGA 540
QY 541 CACCCATCAACAGTATTTATTTTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTATTTTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGCATTTGGGTACACAGCAATCGCGTGTAGCGGGCCCAATTAAAGTTCTGCTCTCGG 660
Db 601 TGGTCGCATTTGGGTACACAGCAATCGCGTGTAGCGGGCCCAATTAAAGTTCTGCTCTCGG 660
QY 661 CGGCTCTGCGTCTGGCTGGCGTAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGGCTCTGCGTCTGGCTGGCGTAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAACGGGAAGCGACTGGAGTGCCTATGTCGGTTCGCGTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGCCTATGTCGGTTCGCGTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGCGATCGTTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGCGATCGTTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGGCATTCACGAGTCCGGCTCGCGTTCGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCATTCACGAGTCCGGCTCGCGTTCGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGTTCGCGGATATCTCGGTAGTGGGATACG 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGTTCGCGGATATCTCGGTAGTGGGATATTC 960
QY 961 GCCTGTGGGGCAAAACCGAGTCCGGCTGTCGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACCGAGTCCGGCTGTCGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAAACCCACTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAAACCCACTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGACAGCTGGGACACAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGACAGCTGGGACACAGAGTTT 1140
QY 1141 CCGCATGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGCATGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGAGCATCGGAAGCTGTGTTATGGCTGTGCAAGGTCTGTAATCACTGCAATAATTCG 1320
Db 1261 TCAGGCGAGCATCGGAAGCTGTGTTATGGCTGTGCAAGGTCTGTAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCGCGTAAATGTTTTTGGCGGACATCAATACCGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCGCGTAAATGTTTTTGGCGGACATCAATACCGTT 1380
QY 1381 CTGGCAAAATATCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTGTGAGCGGATACCAATTTTCAACAGGAAACAGCGAGTCCGTTTGGTGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATACCAATTTTCAACAGGAAACAGCGAGTCCGTTTGGTGTGTTTTCACGA 1500
QY 1501 GCACCTTCACCAACAGGACCATAGATTATGAAGCTGAAGAGGTAAACTGTTGTAATCTCGG 1560
Db 1501 GCACCTTCACCAACAGGACCATAGATTATGAAGCTGAAGAGGTAAACTGTTGTAATCTCGG 1560
QY 1561 ATTAAACGGGATAAAGGCTTAAACCGTCTCGCTGAAGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATAAAGGCTTAAACCGTCTCGCTGAAGTAAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAATTCACAGGTT 1680
Db |||||||
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAATTCACAGGTT 1680
Db |||||||
Qy 1681 GCGGCAATCGCGATGCGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTTGGTGCTAC 1740
Db |||||||
Qy 1681 GCGGCAATCGCGATGCGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTTGGTGCTAC 1740
Db |||||||
Qy 1741 GCTCAATCTGCGCTGTTGGCTGAATATCAACCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db |||||||
Qy 1801 CCGTTTACCTGGATGCGGTAGCTTACAAACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db |||||||
Qy 1801 CCGTTTACCTGGATGCGGTAGCTTACAAACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db |||||||
Qy 1861 GAAGGTTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCCAAAACCTGGGAA 1920
Db |||||||
Qy 1861 GAAGGTTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCCAAAACCTGGGAA 1920
Db |||||||
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db |||||||
Qy 1981 CTGGAAGAACCGCTACTTCACTGCGCGCTGATGCTGTCGACGGGGTTATGCGTTCAAG 2040
Db |||||||
Qy 1981 CTGGAAGAACCGCTACTTCACTGCGCGCTGATGCTGTCGACGGGGTTATGCGTTCAAG 2040
Db |||||||
Qy 2041 TATGAARACGCGAGTACGACATTAAGACGCTGGCGTGGATTAAGCTGGCGCGAAGCG 2100
Db |||||||
Qy 2041 TATGAARACGCGAGTACGACATTAAGACGCTGGCGTGGATTAAGCTGGCGCGAAGCG 2100
Db |||||||
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAACAAACACATGAAATGCAACACCGATTAC 2160
Db |||||||
Qy 2161 TCCATCGAGAAGCTGCTTAAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db |||||||
Qy 2221 GCATGGTCCACATCGACACGAGCAAGTGAATGTTGTTGTAACCGTACTCGCGACCTTC 2280
Db |||||||
Qy 2281 AAGGTCACACATCCAAACCGTTCTGTCGCTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Db |||||||
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db |||||||
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGTAGCGCTGAAGTCTTACGAGAAAG 2460
Db |||||||
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGTAGCGCTGAAGTCTTACGAGAAAG 2460
Db |||||||
Qy 2461 TTGGCGAAGATCCACGTTATGCGCCACCACTGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db |||||||
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db |||||||
Qy 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||||||
Qy 2641 AACACAAACAACAATAACAAACAACTCGGGATCGAGGGAAGGATTTCAGAATTC 2700
Db |||||||
Qy 2641 AACACAAACAACAATAACAAACAACTCGGGATCGAGGGAAGGATTTCAGAATTC 2700
Db |||||||

Qy 2701 GGATCCTCT 2709
Db |||||||
Qy 2701 GGATCCTCT 2709
Db |||||||
RESULT 8
US-10-343-859-8
; Sequence 8, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; TITLE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; PRIOR FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 6648
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
; OTHER INFORMATION: pMAL-c2x
US-10-343-859-8

Query Match 81.8%; Score 2701; DB 19; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCCCGCGAAGAGA 60
Db |||||||
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCCCGCGAAGAGA 60
Db |||||||
Qy 61 GTCAATTCCAGGTTGGTGAATGTGAAACCAAGTAAAGCTTATACGATGTCTCGACAGTATGCG 120
Db |||||||
Qy 61 GTCAATTCCAGGTTGGTGAATGTGAAACCAAGTAAAGCTTATACGATGTCTCGACAGTATGCG 120
Db |||||||
Qy 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGTAACCAAGCCAGCCAGCCACCTTTCTGCGAAAA 180
Db |||||||
Qy 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGTAACCAAGCCAGCCAGCCACCTTTCTGCGAAAA 180
Db |||||||
Qy 181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db |||||||
Qy 181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db |||||||
Qy 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGGCCCTGC 300
Db |||||||
Qy 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGGCCCTGC 300
Db |||||||
Qy 301 ACGCCCGTTCGCAAAATTTGTCGCGCGATTAATCTCGCGCCGATCAATCTGGGTGTCACGCG 360
Db |||||||
Qy 301 ACGCCCGTTCGCAAAATTTGTCGCGCGATTAATCTCGCGCCGATCAATCTGGGTGTCACGCG 360
Db |||||||
Qy 361 TGGTGGTTCGATGGTAGAACGAGCGCGTGAAGCCCTGTAAGCGCGGTGCAACAATC 420
Db |||||||
Qy 361 TGGTGGTTCGATGGTAGAACGAGCGCGTGAAGCCCTGTAAGCGCGGTGCAACAATC 420
Db |||||||
Qy 421 TTCTCGGCAACCGGTGATGGGCTGATCAATTAATCTATCCCTGATCAACAGGATGCCA 480
Db |||||||
Qy 421 TTCTCGGCAACCGGTGATGGGCTGATCAATTAATCTATCCCTGATCAACAGGATGCCA 480
Db |||||||
Qy 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGGCTTATTTCTTGAATCTCTGACCCAGA 540
Db |||||||
Qy 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGGCTTATTTCTTGAATCTCTGACCCAGA 540
Db |||||||
Qy 541 CACCATCAACAGTATTTATTTTCTCCCATGAAGCGGTACGGGACTGGGGCTGGAGCATC 600
Db |||||||

Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGCGCGTGGAGCATC 600
Qy 601 TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCCCAATTAAGTTCTGTCGCG 660
Db 601 TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCCCAATTAAGTTCTGTCGCG 660
Qy 661 CCGGTCGTGGTCTGGCTGGCTGGCAATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGGTCGTGGTCTGGCTGGCTGGCAATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CCGAAACGGGAAGCGCACTGGAGTGCATGTCGGTTTTCACAAACCAATGCAATGCTGA 780
Db 721 CCGAAACGGGAAGCGCACTGGAGTGCATGTCGGTTTTCACAAACCAATGCAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGAA 840
Qy 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACCATCAAAACAGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACCATCAAAACAGATTTTC 960
Qy 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGCGCCCTCTCACTGGTGAAGAAACCAACCTGCGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCCCTCTCACTGGTGAAGAAACCAACCTGCGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGCGTGTGGCGGATTCATTAATGCAGCTGSCACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCGTGTGGCGGATTCATTAATGCAGCTGSCACGACAGGTTT 1140
Qy 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTCTAGCTCACTCAATTAG 1200
Db 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTCTAGCTCACTCAATTAG 1200
Qy 1201 GCACAATTCATGTTTGACAGCTTATCATCGATGCAACGCTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGATGCAACGCTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGTATGGCTGTGACGCTCGTAAATCACTGCAATAATCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGTATGGCTGTGACGCTCGTAAATCACTGCAATAATCG 1320
Qy 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAACCGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAACCGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATCAAAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCACGA 1500
Db 1441 ATTGTAGCGGATCAAAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCACGA 1500
Qy 1501 GCATTCACCAACAGGACCATAGATTAATGAACCTGAAGAAAGTAACTGGTAAATCTGG 1560
Db 1501 GCATTCACCAACAGGACCATAGATTAATGAACCTGAAGAAAGTAACTGGTAAATCTGG 1560
Qy 1561 ATTAACGGCGATTAAGGCTATACGGTCTCGCTGAAGTAACTGGAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGCGATTAAGGCTATACGGTCTCGCTGAAGTAACTGGAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680

Qy 1581 CGGCAAACTGGCGATGCGCCTGACATTAATCTTCTGGGCAACACGACCGCTTTGGTGGCTAC 1740
Db 1581 CGGCAAACTGGCGATGCGCCTGACATTAATCTTCTGGGCAACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCCGTACGTTTACAACGCGCAAGCTGATGTTCTTACC CGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCCGTACGTTTACAACGCGCAAGCTGATGTTCTTACC CGATCGCTGTT 1860
Qy 1861 GAGCGTTATCGCTGATTTTATAAAGAACTGAAAGCGAAGGTTAAGAGCGCGCTGATGTTCAAC 1920
Db 1861 GAGCGTTATCGCTGATTTTATAAAGAACTGAAAGCGAAGGTTAAGAGCGCGCTGATGTTCAAC 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGCGCGCTGATTTGCTGTAACGCGGCTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGCGCGCTGATTTGCTGTAACGCGGCTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGGCAAGTAGACATTTAAAGAGCTGGCGGTGGATTAACCGTGGCGGAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTAGACATTTAAAGAGCTGGCGGTGGATTAACCGTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACATGATGACAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACATGATGACAGACACCGGATTAC 2160
Qy 2161 TCCATCGCAGAACCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAACCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Qy 2221 GCATGTCCAAACATCACACACCGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGTCCAAACATCACACACCGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACACATTCACACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACACATTCACACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTCGCGCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTCGCGCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCGTGGCTGCTGCGTGGTGAATCAACGC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCGTGGCTGCTGCGTGGTGAATCAACGC 2580
Qy 2581 GCCAGCGTCTGACACTGTTCGATGAAGCCTTGAAGAGCGCGACACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTTCGATGAAGCCTTGAAGAGCGCGACACTAAATTCGAGCTCG 2640
Qy 2641 AACACAAACAAATAAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATAAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCTCT 2709
Db 2701 GGATCTCTCT 2709

RESULT 9

US-10-343-859-9
; Sequence 9, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; TITLE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9191
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-muts
; OTHER INFORMATION: Plasmid
US-10-343-859-9

Query Match 81.8%; Score 2698; DB 19; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	CGGACACCATCGAATGGTGCAAAA	CTTTCCGGGTATGGCATGATAGAGCCCGGAAGAGA	60
DB	1	CGGACACCATCGAATGGTGCAAAA	CTTTCCGGGTATGGCATGATAGAGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAAC	CAGTAACTATACCATGTCGAGAGTATCGG	120
DB	61	GTCAATTCAGGGTGGTGAATGTGAAC	CAGTAACTATACCATGTCGAGAGTATCGG	120
QY	121	GTGTCCTTTATCAGACCGTTTCCG	CGGTGGAACAGCCGACGCACTCTCGGAAAA	180
DB	121	GTGTCCTTTATCAGACCGTTTCCG	CGGTGGAACAGCCGACGCACTCTCGGAAAA	180
QY	181	CCGGGAAAAAGTGAAGCGCGATG	GGAGCTGAATTTACATTTCCAAACCGGTGGCAC	240
DB	181	CCGGGAAAAAGTGAAGCGCGATG	GGAGCTGAATTTACATTTCCAAACCGGTGGCAC	240
QY	241	AACAATCGCGGGCAACAGTCTG	TGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAATCGCGGGCAACAGTCTG	TGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGGTGCGCAAAATTTG	CGCGGGATTAATCTCGCGCGATCAACTGGGTGGCAGCG	360
DB	301	ACGCGCGGTGCGCAAAATTTG	CGCGGGATTAATCTCGCGCGATCAACTGGGTGGCAGCG	360
QY	361	TGGTGGTTCGATGTGATAGAAC	GAAGCGGTGTAAGCGCGGTGCAAAATC	420
DB	361	TGGTGGTTCGATGTGATAGAAC	GAAGCGGTGTAAGCGCGGTGCAAAATC	420
QY	421	TTCTCGCGCAACCGTCAGTGG	CGGTGATTAATCTATCCGTGGATGACACAGGATGCCA	480
DB	421	TTCTCGCGCAACCGTCAGTGG	CGGTGATTAATCTATCCGTGGATGACACAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCCTCA	TATATTTCCGGCGTTATTTCTGATGTCCTGACCCAGA	540
DB	481	TTGCTGTGGAAGCTGCCTCA	TATATTTCCGGCGTTATTTCTGATGTCCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTTAT	TTTCTCCATGAAGACGGTACGGGCTGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTTAT	TTTCTCCATGAAGACGGTACGGGCTGGGCGTGGAGCATC	600
QY	601	TGGTGGCATGGGTGACAGAA	ATCGCGCTGTAGCGGGCCCATTTAAGTTCTGTCTCGG	660
DB				

DB	601	TGGTGGCATGGGTGACACAGCA	AAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGCTGGCT	GGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGCTGGCT	GGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGCACTGG	AGTGGCATGTCCGGTTTTTCAACAAACCAATGCAATGCTGA	780
DB	721	CGGAACGGGAAGGCACTGG	AGTGGCATGTCCGGTTTTTCAACAAACCAATGCAATGCTGA	780
QY	781	ATGAGGGCATGTTCCCACT	CGCATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA	840
DB	781	ATGAGGGCATGTTCCCACT	CGCATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA	840
QY	841	TGCGCGCATTAACGAGTCC	CGGCTGCGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGCGCATTAACGAGTCC	CGGCTGCGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCT	CATGTTATATCCGCGCGTTAACCCACATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCT	CATGTTATATCCGCGCGTTAACCCACATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAAAACCA	CGCTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
DB	961	GCCTGCTGGGCAAAACCA	CGCTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
QY	1021	AGGGCAATCAGCTGTGTC	CGCTCTCATGTGTAAGAAAAAACCCCTGGCGCCCAATA	1080
DB	1021	AGGGCAATCAGCTGTGTC	CGCTCTCATGTGTAAGAAAAAACCCCTGGCGCCCAATA	1080
QY	1081	CGAATACCGCTCTCCCG	CGCGTGGCGGATTCATTAATGAGTGTAGTCTACTCATTTAG	1140
DB	1081	CGAATACCGCTCTCCCG	CGCGTGGCGGATTCATTAATGAGTGTAGTCTACTCATTTAG	1140
QY	1141	CCGACTCGGAAAGCGG	CAGTGAGCGCAACCAATTAATGTAGTGTAGTCTACTCATTTAG	1200
DB	1141	CCGACTCGGAAAGCGG	CAGTGAGCGCAACCAATTAATGTAGTGTAGTCTACTCATTTAG	1200
QY	1201	GCACAATTTCTCATGTT	TGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGCG	1260
DB	1201	GCACAATTTCTCATGTT	TGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCATCGGAAG	CTGTGTATGCTGTGAGAGTCTGTAAATCTACTGATTAATTCG	1320
DB	1261	TCAGGACGCATCGGAAG	CTGTGTATGCTGTGAGAGTCTGTAAATCTACTGATTAATTCG	1320
QY	1321	TGTGCTCAAGCGCACT	CCCGTTCTGGAATATGTTTTTGGCGCGACATCAATACGTT	1380
DB	1321	TGTGCTCAAGCGCACT	CCCGTTCTGGAATATGTTTTTGGCGCGACATCAATACGTT	1380
QY	1381	CTGGCAAAATTTCTGAA	ATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440
DB	1381	CTGGCAAAATTTCTGAA	ATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440
QY	1441	ATTGTGAGCGGATAACA	ATTTCAACAGGAACCGCAGTCCGTTAGTGTGTTTTCACGA	1500
DB	1441	ATTGTGAGCGGATAACA	ATTTCAACAGGAACCGCAGTCCGTTAGTGTGTTTTCACGA	1500
QY	1501	GCACCTTCAACCAAGG	ACCATAGATTAATACTGAAGAAGGTAAATCTGGTAATCTGG	1560
DB	1501	GCACCTTCAACCAAGG	ACCATAGATTAATACTGAAGAAGGTAAATCTGGTAATCTGG	1560
QY	1561	ATTAACGGCGATTAAG	GGCTATAACCGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAACGGCGATTAAG	GGCTATAACCGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCA	CCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCACAGTT	1680
DB	1621	ACCGGAATTAAGTCA	CCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCACAGTT	1680
QY	1681	GGGGCAATCGGGGAT	GGCGCTGACATTAATCTTCTGGGCACACGCCCTTTGGTGGCTAC	1740
DB	1681	GGGGCAATCGGGGAT	GGCGCTGACATTAATCTTCTGGGCACACGCCCTTTGGTGGCTAC	1740

```
QY 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCCGTAGCTTTACAAAGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCCGTAGCTTTACAAAGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCGCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCGCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGAGCAAGGTAAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGAGCAAGGTAAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAAGCGCAAGTACGACATTAAGACAGCTGGGCGTGGATTAAGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAAAGCGCAAGTACGACATTAAGACAGCTGGGCGTGGATTAAGCTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGATTAAGAACTGAATTAAGAAACCAACACATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGATTAAGAACTGAATTAAGAAACCAACACATGAATGACACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTAATAAGCGCAACAGCGATGACCATCAACGGCGCGGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAGCGCAACAGCGATGACCATCAACGGCGCGGTGG 2220
QY 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAAGTGTAAAGCTGCTGCCAGCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAAGTGTAAAGCTGCTGCCAGCTTC 2280
QY 2281 AAGGTCAACCATCAACCGTTGCTGGCTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGTCAACCATCAACCGTTGCTGGCTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCCAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCCAAACCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAAGTCTTACGAGGAAAG 2460
Db 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAAGTCTTACGAGGAAAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTTGCCGCAACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTTGCCGCAACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCCGCTGGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCCGCTGGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTGTGAGATGTCGCTTTCTGGTATCCGCTGGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 2581 GCCAGCGTGTGAGATGTCGCTTTCTGGTATCCGCTGGCTGCTGCTGCTGCTGCTGCTGCTG 2640
QY 2641 AACCAACAAACATTAACATTAACAAACCTCGGGATCGAGGGAAGGATTTGAGATTC 2700
Db 2641 AACCAACAAACATTAACATTAACAAACCTCGGGATCGAGGGAAGGATTTGAGATTC 2700
QY 2701 GGATCC 2706
Db 2701 GGATCC 2706
```

RESULT 10
US-10-263-153-40
; Sequence 40, Application US/10263153

```
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 7112
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3189)
; OTHER INFORMATION: pMBP-c2K-Toxop30del11 (52-214aa)
; US-10-263-153-40
```

```
Query Match 81.5% Score 2690.4; DB 18; Length 7112;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCAATTACGGGTGGTGAATGTGAACCCAGTACGTTATACGATGTCGAGATATGCCG 120
Db 61 GTCAATTACGGGTGGTGAATGTGAACCCAGTACGTTATACGATGTCGAGATATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCCAGCTTCTTCCGAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCCAGCTTCTTCCGAAA 180
QY 181 CCGCGGAAAAGTGGAGCGCGCATGGCGAGCTGAATTACATTCGCAACCGCGTGGCAC 240
Db 181 CCGCGGAAAAGTGGAGCGCGCATGGCGAGCTGAATTACATTCGCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACACAGTCTGCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACACAGTCTGCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTGC AAAATTGTGCGCGCATTAATAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTGC AAAATTGTGCGCGCATTAATAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGCGATGGTAGAACGAGCGGCTGAGACCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGGTGTGTGCGATGGTAGAACGAGCGGCTGAGACCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGGTCAGTGGCTGATCAATTAATATCCGCTGGATGACCGAGATGCCA 480
Db 421 TTCTCGCGCAACCGGTCAGTGGCTGATCAATTAATATCCGCTGGATGACCGAGATGCCA 480
QY 481 TTGCTGTGAAGCTGCTGCACCTAATTTTCGCGGTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACCTAATTTTCGCGGTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTGTGATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGTCTCGG 660
Db 601 TGGTGTGATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGTCTCGG 660
QY 661 CCGGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
```

[illegible]

RESULT 11
US-10-263-153-30
; Sequence 30, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.

/ TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
/ FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
/ CURRENT APPLICATION NUMBER: US/10/263,153
/ CURRENT FILING DATE: 2002-10-22
/ NUMBER OF SEQ ID NOS: 74
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 30
/ LENGTH: 7259
/ TYPE: DNA
/ ORGANISM: Toxoplasma gondii
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1528)...(3336)
/ OTHER INFORMATION: pMBP-c2X-Toxop30del14del18 (83-294aa)
US-10-263-153-30

Query Match 81.5%; Score 2690.4; DB 18; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCGGAGAGA 60

QY 61 GTCGAATTCAGGGTGTGAATGTGAACCCAGTAACGTTATACGATGTTCGACAGTATGCCG 120
Db 61 GTCGAATTCAGGGTGTGAATGTGAACCCAGTAACGTTATACGATGTTCGACAGTATGCCG 120

QY 121 GTGCTCTTATCAGACCGTGTTCGCGGTGGTGAACCGCCAGCCACGTTTCTCGGAAA 180
Db 121 GTGCTCTTATCAGACCGTGTTCGCGGTGGTGAACCGCCAGCCACGTTTCTCGGAAA 180

QY 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTAATCCCAACCGCGTGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTAATCCCAACCGCGTGCAC 240

QY 241 AACAACTGGCGGCAAAACAGTGTGTGTGAATGGCGTGTGCACCTCCAGTCTGCGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTGTGTGTGAATGGCGTGTGCACCTCCAGTCTGCGCCCTGC 300

QY 301 ACGGCGCTGCGCAATGTGCGCGGATTAATCTCGCGCGATCACTGGTGGCCAGCG 360
Db 301 ACGGCGCTGCGCAATGTGCGCGGATTAATCTCGCGCGATCACTGGTGGCCAGCG 360

QY 361 TGGTGGTGTGATGGTGAACGAAGCGCGTGAAGCGCTGAAGCGCGGTGCAATC 420
Db 361 TGGTGGTGTGATGGTGAACGAAGCGCGTGAAGCGCTGAAGCGCGGTGCAATC 420

QY 421 TTCTCGGCAACCGGTGATGGGTGATCAATTAATCTATCGCTGGATGAACGAGATGCCA 480
Db 421 TTCTCGGCAACCGGTGATGGGTGATCAATTAATCTATCGCTGGATGAACGAGATGCCA 480

QY 481 TTGCTGTGGAGCTGCTGCACTAATGTTCGCGGTATTTCTTGATGTCTCTGACCGA 540
Db 481 TTGCTGTGGAGCTGCTGCACTAATGTTCGCGGTATTTCTTGATGTCTCTGACCGA 540

QY 541 CACCCATCAACAGTATTTATTTCTCCATGAAGACGCTACCGGCTGGGCTGGAGATC 600
Db 541 CACCCATCAACAGTATTTATTTCTCCATGAAGACGCTACCGGCTGGGCTGGAGATC 600

QY 601 TGGTGGATTTGGGTCAACAGCAATTCGCTGTAGCGGCGCATTAAGTTCGTCTCGG 660
Db 601 TGGTGGATTTGGGTCAACAGCAATTCGCTGTAGCGGCGCATTAAGTTCGTCTCGG 660

QY 661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAAACGGGAGCGCATGGAGTGCATGTCCGGTTTCAACAAACCATCAATGCTGA 780
Db 721 CGGAAACGGGAGCGCATGGAGTGCATGTCCGGTTTCAACAAACCATCAATGCTGA 780

QY 781 ATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGCGCTGGCGCAA 840
Db 781 ATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGCGCTGGCGCAA 840

QY 841 TGGCGGCATATTACCGAGTCCGGGCTGCGGTGGTGGGATATCTCGGTAGTGGATACG 900
Db 841 TGGCGGCATATTACCGAGTCCGGGCTGCGGTGGTGGGATATCTCGGTAGTGGATACG 900

QY 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAAAACAGATTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAAAACAGATTTTC 960

QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTCTCACTGGTGAAGAAAAACCCCTCGCGCCCAATA 1080
Db 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTCTCACTGGTGAAGAAAAACCCCTCGCGCCCAATA 1080

QY 1081 CGMAACCGGCTCTCCCGCGGTGGCCGATTCATTAATGACAGTGGCAGCAGAGTTT 1140
Db 1081 CGMAACCGGCTCTCCCGCGGTGGCCGATTCATTAATGACAGTGGCAGCAGAGTTT 1140

QY 1141 CCGACTGGAAAGCGGSCAGTGAACGCAACGCAATTAATGAGTGGTGGTGGTGGTGGTGG 1200
Db 1141 CCGACTGGAAAGCGGSCAGTGAACGCAACGCAATTAATGAGTGGTGGTGGTGGTGGTGG 1200

QY 1201 GCACAATTTCTCATGTTTGAAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGAAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCG 1260

QY 1261 TCAGGAGCCATCGGAAGCTGCTGATGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGCTGATGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1320

QY 1321 TGTGCTCAAGCGGCACTCCCGTTCCTGGAATAATGTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCCTGGAATAATGTTTTTGGCGCGACATCAACGGTT 1380

QY 1381 CTGGCAATTTCTGAAATGAGCTGTGACAAATTAATCATCGCTCGTAAATGTTGTTGGA 1440
Db 1381 CTGGCAATTTCTGAAATGAGCTGTGACAAATTAATCATCGCTCGTAAATGTTGTTGGA 1440

QY 1441 ATTGTGAGCGGATAACAAATTTACAGAGGAAACAGCGCTGGTGGTGGTGGTGGTGG 1500
Db 1441 ATTGTGAGCGGATAACAAATTTACAGAGGAAACAGCGCTGGTGGTGGTGGTGGTGG 1500

QY 1501 GCACTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAACCTGGTAAATCTGG 1560
Db 1501 GCACTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAACCTGGTAAATCTGG 1560

QY 1561 ATTAAACGGGATTAAGGCTTAACCGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATTAAGGCTTAACCGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCAACCGTGGAGTCCGATTAACCTGGAAGAGAAATTCGAGAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTGGAGTCCGATTAACCTGGAAGAGAAATTCGAGAGGTT 1680

QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGACACAGGACCGCTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGACACAGGACCGCTTGGTGGCTAC 1740

QY 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGAACAAAGCGTTCCAGAACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGAACAAAGCGTTCCAGAACAGCTGTAT 1800

QY 1801 CGGTTTACCTGGGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Db 1801 CGGTTTACCTGGGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860

QY 1861 GAAGCGTTATCGCTGATTTATTAACAAAGATCTGCTGCGCAACCCCGCAAAACCTGGGAA 1920


```
Db 1861 GAAGGTTATCGCTGATTTATACAAAGATCGTCGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCGGCGCTGGATAAAGAACTGAAAGCGAAGGTAGAGCGCGCTGATGTTCAAC 1980
Db 1971 GAGATCCGGCGCTGGATAAAGAACTGAAAGCGAAGGTAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTCGGCGCTGATGCTGTGACGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTCGGCGCTGATGCTGTGACGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAARACGCACTAGCACATTAAGACGCTGGCGTGGATAACGCTGGCGGAAGCG 2100
Db 2041 TATGAARACGCACTAGCACATTAAGACGCTGGCGTGGATAACGCTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAAGCGAAGGTAGATGTTCAACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAAGCGAAGGTAGATGTTCAACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGAAGGTAGATGTTCAACACCGATTAC 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGAAGGTAGATGTTCAACACCGATTAC 2220
Qy 2221 GCATGGTCAACATCGACACCAAGAGTGAATATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACCAAGAGTGAATATGTTGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCAATCAACACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACCGCCCGCAGT 2340
Db 2281 AAGGTCACCAATCAACACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACCGCCCGCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTATTAAGACAAACCGCTGGTGGCGTGTGAGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTATTAAGACAAACCGCTGGTGGCGTGTGAGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAGATCCACGATTTGCCGCCACCACTGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGATTTGCCGCCACCACTGGAAGAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGTTGATGCGGTGATGCGGTGATCAACGCC 2580
Db 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGTTGATGCGGTGATGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGGAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGGAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACCAACAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACA 2700
Db 2641 AACCAACAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACA 2700
```

RESULT 12

```
US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US2004006729A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 35
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3399)
; OTHER INFORMATION: pMBP-c2X-Toxop30del10 (52-284aa)
US-10-263-153-35
```

Query Match

```
Best Local Similarity 81.5%; Score 2690.4; DB 18; Length 7322;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGCGTATGCGCATGATAGCCCGGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGCGTATGCGCATGATAGCCCGGGAAGAGA 60
Qy 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGTTATCGATGTCGAGATGCGG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGTTATCGATGTCGAGATGCGG 120
Qy 121 GTGTCCTTATCAGACCGCTTTCCCGCGTGTGAACCCAGGCGAGCCAGCCCTTTCTGCGAAAA 180
Db 121 GTGTCCTTATCAGACCGCTTTCCCGCGTGTGAACCCAGGCGAGCCAGCCCTTTCTGCGAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGCATGCGCGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGCATGCGCGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGGCAACAGTCTGTTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGGCAACAGTCTGTTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTTCGCAAAATTCGTCGCGCGATTAATATCTCGCGCGCATCAATCGGTGGTGCAGG 360
Db 301 ACGCGCGTTCGCAAAATTCGTCGCGCGATTAATATCTCGCGCGCATCAATCGGTGGTGCAGG 360
Qy 361 TGGTGGTTCGATGATAGAACGAGCGCGTGAAGCGCTGAAGCGCGGTGCAAAATC 420
Db 361 TGGTGGTTCGATGATAGAACGAGCGCGTGAAGCGCTGAAGCGCGGTGCAAAATC 420
Qy 421 TTCTCGCGCAACCGGTGATGCGGTGATCAATTAATCTCGGTGATGATGATGATGATGATGATG 480
Db 421 TTCTCGCGCAACCGGTGATGCGGTGATCAATTAATCTCGGTGATGATGATGATGATGATGATG 480
Qy 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGCGCTTATTTCTGATGTTCTGATGATGATGATG 540
Db 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGCGCTTATTTCTGATGTTCTGATGATGATGATG 540
Qy 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGCGGTACGCGACTGGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGCGGTACGCGACTGGGCGGTGGAGCATC 600
Qy 601 TGGTGCATGTTGGTCAACAGCAAAATCGCGTGTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGCATGTTGGTCAACAGCAAAATCGCGTGTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGGCTGGCTGGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTGGCTGGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTGTGCGAAGCATGATGCGGTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTGTGCGAAGCATGATGCGGTGGCGGCA 840
Qy 841 TCGCGGCCATTACCGAGTCCGGGCTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
```

```
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCGCGCTCTCACTGCTGAAAGAAACACACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCGCGCTCTCACTGCTGAAAGAAACACACCTGGCGGCCAATA 1080
QY 1081 CCGAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCACTGGCAGCACAGGTTT 1140
Db 1081 CCGAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCACTGGCAGCACAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCAATTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCAATTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGTTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGTTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTCTCTGGAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTCTCTGGAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1440
QY 1441 ATTGTGAGCGGATAACATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500
QY 1501 GCATTTTCAACCAACAGGACCATAGATTATGAAACTGAAAGAGTAAACTGGTAACTCTGG 1560
Db 1501 GCATTTTCAACCAACAGGACCATAGATTATGAAACTGAAAGAGTAAACTGGTAACTCTGG 1560
QY 1561 ATTTAAGCGCGATAAGGCTATAACCGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGCGATAAGGCTATAACCGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCACCGTTGAGCATTCCGATTAACCTGGNAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGAGCATTCCGATTAACCTGGNAGAAATTTCCACAGGTT 1680
QY 1681 CGCGCAACTGCGGATGCGCCCTGACATATTCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGCGGATGCGCCCTGACATATTCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTTGGCTGAAATCAACCCCGACAAAGGTTTCCAGGCAACAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTTGGCTGAAATCAACCCCGACAAAGGTTTCCAGGCAACAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTTACACGGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTTACACGGCAAGCTGATGCTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGTGTATTATAACAAAGATCTGCTGCCGAACCGCCCAAAACCTGGAA 1920
Db 1861 GAAGCGTTATCGTGTATTATAACAAAGATCTGCTGCCGAACCGCCCAAAACCTGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGACTGAAAGCGAAGGTAAGAGCGGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGACTGAAAGCGAAGGTAAGAGCGGCTGATGTTCAAC 1980
```

```
QY 1981 CTGCAAGAACGCTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTAGTACGACATTAAAGAGCGTGGCGCTGGATACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTAGTACGACATTAAAGAGCGTGGCGCTGGATACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAAAAAAACAACATGATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAAAAAAACAACATGATGACAGACACCGATTAC 2160
QY 2161 TCCATCCGAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCGCGGTGG 2220
Db 2161 TCCATCCGAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCGCGGTGG 2220
QY 2221 GCATGTTCCAAACATCGACACCGCAAGTGAATTTATGGTGTAAACGGTACTTGCAGACCTTC 2280
Db 2221 GCATGTTCCAAACATCGACACCGCAAGTGAATTTATGGTGTAAACGGTACTTGCAGACCTTC 2280
QY 2281 AAGGGTCAACCATCCCAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGGCGCGCAGT 2340
Db 2281 AAGGGTCAACCATCCCAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGGCGCGCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTTAAATAAGCAAAACCGTGGTGGCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTTAAATAAGCAAAACCGTGGTGGCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGCGGAAAGATCCAGTATTGCGCCACCATCGCAAAACCGTGGTGGCTGAGCGTGAAGTCTT 2520
Db 2461 TTGCGGAAAGATCCAGTATTGCGCCACCATCGCAAAACCGTGGTGGCTGAGCGTGAAGTCTT 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTCGGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTCGGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGCTGAGACTGCTGATGAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGCTGAGACTGCTGATGAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAAACATAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2700
Db 2641 AACAAACAAACAAATAAACATAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2700
```

RESULT 13

```
US-10-263-153-25
; Sequence 25, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 7352
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3429)
```


OTHER INFORMATION: pMBP-c2X-Toxop30del14C (52-294aa)
US-10-263-153-25

Query Match 81.5%; Score 2690.4; DB 18; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA	60
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA	60
Qy	61	GTCAATTCCAGGCTGTGAATGTGAAGAACAGTAAGCTTATACGATGCTCCAGAGTATGCCG	120
Db	61	GTCAATTCCAGGCTGTGAATGTGAAGAACAGTAAGCTTATACGATGCTCCAGAGTATGCCG	120
Qy	121	GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACAGCGCCAGCCACCGTTTCTCGAAAA	180
Db	121	GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACAGCGCCAGCCACCGTTTCTCGAAAA	180
Qy	181	CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC	240
Db	181	CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC	240
Qy	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Qy	301	ACGGCCGCTCGCAAAATTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
Db	301	ACGGCCGCTCGCAAAATTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
Qy	361	TGGTGGTTCGATGGTAGAAAGAAAGCGCGTGAAGCGCTGAAAGCGCGGTGCAACAATC	420
Db	361	TGGTGGTTCGATGGTAGAAAGAAAGCGCGTGAAGCGCTGAAAGCGCGGTGCAACAATC	420
Qy	421	TTCTCGCCCAACGGTCAGTGGGTGATCAATTAATCTCCGTGGATCAACAGATGCCA	480
Db	421	TTCTCGCCCAACGGTCAGTGGGTGATCAATTAATCTCCGTGGATCAACAGATGCCA	480
Qy	481	TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCA	540
Db	481	TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCA	540
Qy	541	CACCCATCAACAGTATTTTCTCCATGAAGCGGTACGCGATCGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTTTCTCCATGAAGCGGTACGCGATCGCGGTGGAGCATC	600
Qy	601	TGGTCCGATGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCGTCTCG	660
Db	601	TGGTCCGATGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCGTCTCG	660
Qy	661	CGCGTCTCGGCTGGCTGGCAATAATCTCACCTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTCGGCTGGCTGGCAATAATCTCACCTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAAACGGAAGCGGACCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAAACGGAAGCGGACCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780
Qy	781	ATGAGGGATCGTTCCCATCTGGATGCTGGTTGCCAAGATCAGATGGCGTGGCGCA	840
Db	781	ATGAGGGATCGTTCCCATCTGGATGCTGGTTGCCAAGATCAGATGGCGTGGCGCA	840
Qy	841	TGCGGCCATTAACAGGTCGGGCTGGGTGGCGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGGCCATTAACAGGTCGGGCTGGGTGGCGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC	960
Db	901	ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC	960
Qy	961	GCCTGCTGGGCAAAACAGAGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020
Db	961	GCCTGCTGGGCAAAACAGAGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020

Db	961	GCCTGCTGGGCAAAACAGAGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020
Qy	1021	AGGCAATACAGCTGTGTCCTGCTCTACTGTGTAAGAAAAACCACTGCGGCCCAATA	1080
Db	1021	AGGCAATACAGCTGTGTCCTGCTCTACTGTGTAAGAAAAACCACTGCGGCCCAATA	1080
Qy	1081	CGGAAACGGCTCTCCCGCGGCTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140
Db	1081	CGGAAACGGCTCTCCCGCGGCTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140
Qy	1141	CCCAGCTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG	1200
Db	1141	CCCAGCTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG	1200
Qy	1201	GCAAAATTCATGTTTGACAGCTTATCATCGATCGATCGATCGATCGATCGATCGATCG	1260
Db	1201	GCAAAATTCATGTTTGACAGCTTATCATCGATCGATCGATCGATCGATCGATCGATCG	1260
Qy	1261	TCAGGACGCAATCGGAAGCTGTGTATGGCTGTGAGGTCTGTAATCACTGCATTAATTCG	1320
Db	1261	TCAGGACGCAATCGGAAGCTGTGTATGGCTGTGAGGTCTGTAATCACTGCATTAATTCG	1320
Qy	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATTAATGTTTTTTCGCGCGCATCATCAACGGTT	1380
Db	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATTAATGTTTTTTCGCGCGCATCATCAACGGTT	1380
Qy	1381	CTGCAAAATTTCTGAAATGAGTGTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
Db	1381	CTGCAAAATTTCTGAAATGAGTGTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
Qy	1441	ATTGTAGCGGATTAACAAATTTTCAACAGGAAACGCCAGTCCGTTTAGGTTTTCACGA	1500
Db	1441	ATTGTAGCGGATTAACAAATTTTCAACAGGAAACGCCAGTCCGTTTAGGTTTTCACGA	1500
Qy	1501	GCCTTTCACCAACAAAGGACCATGATTAATGAAAACTGAAGAGGTAAATCTGTAATCTCG	1560
Db	1501	GCCTTTCACCAACAAAGGACCATGATTAATGAAAACTGAAGAGGTAAATCTGTAATCTCG	1560
Qy	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGAATTAAGTCAACGGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT	1680
Db	1621	ACCGAATTAAGTCAACGGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT	1680
Qy	1681	CGGCACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTTGGTGGCTAC	1740
Db	1681	CGGCACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCTTTGGCTGAAATCACCCCGCAACAAAGGTTTCAGGACAAAGCTGAT	1800
Db	1741	GCTCAATCTGGCCTTTGGCTGAAATCACCCCGCAACAAAGGTTTCAGGACAAAGCTGAT	1800
Qy	1801	CCGTTTACCTGGGATGGCGTACGTTAACAACGCAAGCTGATGTTTACCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGGCGTACGTTAACAACGCAAGCTGATGTTTACCGATCGCTGTT	1860
Qy	1861	GAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTCACTGCGCGTGAATGCTGCTGACGGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGCGCGTGAATGCTGCTGACGGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAACCGCAAGTACGATTAAGACCGTGGCGGTGAATACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAACCGCAAGTACGATTAAGACCGTGGCGGTGAATACGCTGGCGCGAAAGCG	2100

```
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db |||||
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db |||||
QY 2161 TCCATCGCAGAGCTCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db |||||
QY 2161 TCCATCGCAGAGCTCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db |||||
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2281 AAGGTCACACCATCCAAACCGTTTCGTTGGCTGCTGAGCGCAGGTATTAAACGCCCCAGT 2340
Db |||||
QY 2281 AAGGTCACACCATCCAAACCGTTTCGTTGGCTGCTGAGCGCAGGTATTAAACGCCCCAGT 2340
Db |||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
Db |||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
Db |||||
QY 2401 GAAGCGGTTAATAAAGACAAACCGTTCGCCACCATGGAAGAAACGCCAGAAAGTGAATCATG 2520
Db |||||
QY 2401 GAAGCGGTTAATAAAGACAAACCGTTCGCCACCATGGAAGAAACGCCAGAAAGTGAATCATG 2520
Db |||||
QY 2461 TTGGCGAAAGATCCACGTTATGCGCCACCATGGAAGAAACGCCAGAAAGTGAATCATG 2580
Db |||||
QY 2461 TTGGCGAAAGATCCACGTTATGCGCCACCATGGAAGAAACGCCAGAAAGTGAATCATG 2580
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCGTTGATGCGGTAGCGGTGATGCGGTGATCAACGCC 2580
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCGTTGATGCGGTAGCGGTGATGCGGTGATCAACGCC 2580
Db |||||
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||||
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||||
QY 2641 AACACACACACATTAACATTAACACACCTCGGATCGAGGAAAGATTCAGAAATTC 2700
Db |||||
QY 2641 AACACACACACATTAACATTAACACACCTCGGATCGAGGAAAGATTCAGAAATTC 2700
Db |||||

RESULT 14
US-10-263-153-20
; Sequence 20, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxo30del13C (52-300aa)
US-10-263-153-20

Query Match 81.5%; Score 2690,4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 CCGACACCATCGAATGCGTGAACAAACCTTTTCGCGGTATGGCATGATAGCGCCGGAAGAGA 60
Db |||||
QY 1 CCGACACCATCGAATGCGTGAACAAACCTTTTCGCGGTATGGCATGATAGCGCCGGAAGAGA 60
Db |||||
QY 61 GTCGAATTCAGGGTGGTGAATGCGTGAACAGTGAACGTTATACGATGTCGAGAGTATGCGG 120
Db |||||
QY 61 GTCGAATTCAGGGTGGTGAATGCGTGAACAGTGAACGTTATACGATGTCGAGAGTATGCGG 120
Db |||||
QY 121 GTGCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCGCAGCCAGTTCTTCGCGAAA 180
Db |||||
QY 121 GTGCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCGCAGCCAGTTCTTCGCGAAA 180
Db |||||
QY 181 CGCGGGAAGAGTGAAGCGCGGATGCGCGAGCTGAATTACATTCACACCGCGTGGCAC 240
Db |||||
QY 181 CGCGGGAAGAGTGAAGCGCGGATGCGCGAGCTGAATTACATTCACACCGCGTGGCAC 240
Db |||||
QY 241 AACAACTGGCGGGCAACAGTTCGTTGCTGATTTCGGCGTTCGCGCTCCAGTCTGGCCCTGC 300
Db |||||
QY 241 AACAACTGGCGGGCAACAGTTCGTTGCTGATTTCGGCGTTCGCGCTCCAGTCTGGCCCTGC 300
Db |||||
QY 301 ACGCGCGTTCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db |||||
QY 301 ACGCGCGTTCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db |||||
QY 361 TGGTGGTTCGATGCTAGAACGAGCGGCTCGAAGCTGTAAAGCGCGCGGTGCACAATC 420
Db |||||
QY 361 TGGTGGTTCGATGCTAGAACGAGCGGCTCGAAGCTGTAAAGCGCGCGGTGCACAATC 420
Db |||||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCGCTGGATGACCGAGATGCCA 480
Db |||||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCGCTGGATGACCGAGATGCCA 480
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTTCGCGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTTCGCGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db |||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGCTGGAGCATC 600
Db |||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGCTGGAGCATC 600
Db |||||
QY 601 TGGTGCATTTGGTCAACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db |||||
QY 601 TGGTGCATTTGGTCAACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db |||||
QY 661 CGGCTCTGCTGCTGGCTGGCTGCGCATTAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
Db |||||
QY 661 CGGCTCTGCTGCTGGCTGGCTGCGCATTAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
Db |||||
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db |||||
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db |||||
QY 781 ATGAGGGCATTCGTTCCCATCGGATGCTGGTTGCCAAGATGCGCTGGCGCGCAA 840
Db |||||
QY 781 ATGAGGGCATTCGTTCCCATCGGATGCTGGTTGCCAAGATGCGCTGGCGCGCAA 840
Db |||||
QY 841 TGGCGCGCATTAACCGAGTCCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db |||||
QY 841 TGGCGCGCATTAACCGAGTCCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
Db |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
Db |||||
QY 961 GCCTGCTGGGCAACACGAGTGGAGCGGTTGCTGCACTCTCTCAGGGCCAGCGCGTGA 1020
Db |||||
QY 961 GCCTGCTGGGCAACACGAGTGGAGCGGTTGCTGCACTCTCTCAGGGCCAGCGCGTGA 1020
Db |||||
QY 1021 AGGGCAATCAGTGTGTCGCTCTCACTGGTGAAGAAACCAACCTCTCAGGGCCAGTATA 1080
Db |||||
QY 1021 AGGGCAATCAGTGTGTCGCTCTCACTGGTGAAGAAACCAACCTCTCAGGGCCAGTATA 1080
Db |||||
QY 1081 CGCAACCGGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGCAGCAGACAGGTTT 1140
Db |||||
```

```
Db 1081 |||||CGAACCCTCTCCCGCGGTTGGCGATTCAATATGCGAGTGGCAGCAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACGTGACGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACGTGACGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCATCGGAAGCTGTGTTATGCTGTGAGGTGCTGTAATCATCTGCATAAATTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGTGTTATGCTGTGAGGTGCTGTAATCATCTGCATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACCTCCCGTCTCGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTCTCGATATGTTTTTGGCGGACATCATACGGTT 1380
Qy 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTGTAATGTGTGA 1440
Db 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTGTAATGTGTGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTCAGCA 1500
Db 1441 ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTCAGCA 1500
Qy 1501 GCATTTCACCAAGGACCATAGATTATGAATACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCATTTCACCAAGGACCATAGATTATGAATACTGAAGAGGTAAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAGGCTATAACGGTCTCGTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAGGCTATAACGGTCTCGTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCACCGTTGAGCATCGGATAAATCGGAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGAGCATCGGATAAATCGGAGAGAAATTCACAGGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATAATCAACCCGGAACAAAGCTTCCAGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATAATCAACCCGGAACAAAGCTTCCAGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCCGTAGCTTACACCGCAAGCTGATTGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCCGTAGCTTACACCGCAAGCTGATTGCTTACCGGATCGCTGTT 1860
Qy 1861 GAAAGCTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAAGCTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCTCGGATAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTCGGATAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAACGTTGGCGGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGTTGGCGGTGGATAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACATGAATTCAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACATGAATTCAGACACCGGATTAC 2160
Qy 2161 TCCATCGAAGAGCTGCCTTTAATTAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGAAGAGCTGCCTTTAATTAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
```

```
Db 2161 TCCATCGAAGAGCTGCCTTTAATTAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAATGGTGAACGGTACTCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAATGGTGAACGGTACTCCGACCTTC 2280
Qy 2281 AAGGGTCNAACCATCAAAACCGTTTGGTGGCGTGTGAGCGCAGGTATTTAAACGCCGCCAGT 2340
Db 2281 AAGGGTCNAACCATCAAAACCGTTTGGTGGCGTGTGAGCGCAGGTATTTAAACGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Qy 2401 GAAAGCGGTTAATAAGACAAACCCGCTGGTCCGCTAGCGCTGAAAGTCTTACGAGAAAG 2460
Db 2401 GAAAGCGGTTAATAAGACAAACCCGCTGGTCCGCTAGCGCTGAAAGTCTTACGAGAAAG 2460
Qy 2461 TTGGGGAAGATCCACGTTATTGGCCGACCATGGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGGAAGATCCACGTTATTGGCCGACCATGGAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGGTCGTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCGTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACACACACATACAAATAACAAACCTCGGGATCGAGGGAAGATTCAGAAATTC 2700
Db 2641 AACACACACACATACAAATAACAAACCTCGGGATCGAGGGAAGATTCAGAAATTC 2700

RESULT 15
US-10-263-153-61
; Sequence 61, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxop30MIX1
US-10-263-153-61

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCCCGCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCCCGCGGAAGAGA 60
Qy 61 GTCATTCAGGGTGGTGAATGTAACCAAGTAAAGCTTATACGATGTCGACAGTATGCCG 120
Db 61 GTCATTCAGGGTGGTGAATGTAACCAAGTAAAGCTTATACGATGTCGACAGTATGCCG 120
```

QY 121 GTGTCTCTTATCAGACCGTTTCCCGCTGGTGAACACGCGCAGCCAGCTTCTCGGAAA 180
DB 121 GTGTCTCTTATCAGACCGTTTCCCGCTGGTGAACACGCGCAGCCAGCTTCTCGGAAA 180
QY 181 CCGCGGAAAAGTGAAGCGGGATGGCGAGCTGAATTAATTTCCAAACCGCGTGCAC 240
DB 181 CCGCGGAAAAGTGAAGCGGGATGGCGAGCTGAATTAATTTCCAAACCGCGTGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGTGTGATGGTGGTGGCAGCTCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTCTGTGTGATGGTGGTGGCAGCTCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCAAATTTGTCCGCGGATTAATCTCGCGCGATCAACTCGGTGGCCAGCG 360
DB 301 ACGCGCGTCCAAATTTGTCCGCGGATTAATCTCGCGCGATCAACTCGGTGGCCAGCG 360
QY 361 TGGTGGTGTGATGGTGAACGAGCGGCTCGAAGCTGTAAAGCGGCGGTGCAAAATC 420
DB 361 TGGTGGTGTGATGGTGAACGAGCGGCTCGAAGCTGTAAAGCGGCGGTGCAAAATC 420
QY 421 TTCTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA 480
DB 421 TTCTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAGCTGGCTGACATAATTTCCCGCGTTATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGGAGCTGGCTGACATAATTTCCCGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGATCGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGATCGGCGTGGAGCATC 600
QY 601 TGGTGCATTTGGTTCACCAACCAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTGCATTTGGTTCACCAACCAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
QY 661 CCGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCGGATAG 720
DB 661 CCGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCGGATAG 720
QY 721 CCGAACGGGAAGGCACTGGAGTGCATGTCCGTTTTTCAACCAACCAATGCTGA 780
DB 721 CCGAACGGGAAGGCACTGGAGTGCATGTCCGTTTTTCAACCAACCAATGCTGA 780
QY 781 ATGAGGGCATCTGTCCCACTGGATGCTGGTGGCCAGATCAGATGGCGCTGGCGCAA 840
DB 781 ATGAGGGCATCTGTCCCACTGGATGCTGGTGGCCAGATCAGATGGCGCTGGCGCAA 840
QY 841 TCGCGCCATTACCGAGTCCGGCTGCGCTGGTGGGATATCTCGGTAGTGGGATACG 900
DB 841 TCGCGCCATTACCGAGTCCGGCTGCGCTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAATCAACAGGATTTTC 960
QY 961 GCCTGTGGGCAAAACAGGTGACCGCTTGTGTCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 961 GCCTGTGGGCAAAACAGGTGACCGCTTGTGTCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTCCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTCCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGMAACCGGCTCTCCCGCGTGGCCGATTCAATTAATGAGCTGGCAGCAGAGGTTT 1140
DB 1081 CGMAACCGGCTCTCCCGCGTGGCCGATTCAATTAATGAGCTGGCAGCAGAGGTTT 1140
QY 1141 CCGACTTGAAGCGGCGTGGAGCGCAACGCAATTAATGTGAGTGTCTCACTATTAG 1200
DB 1141 CCGACTTGAAGCGGCGTGGAGCGCAACGCAATTAATGTGAGTGTCTCACTATTAG 1200

QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGGTGACCAATGCTTCGGC 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGGTGACCAATGCTTCGGC 1260
QY 1261 TCAGGAGCAATCGGAAGCTGTGGTATGGCTGTCAGGTCTGTAATCACTGCATAATTCG 1320
DB 1261 TCAGGAGCAATCGGAAGCTGTGGTATGGCTGTCAGGTCTGTAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
DB 1321 TGTGCTCAAGCGCACCTCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTGTAATGTTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTGTAATGTTGGA 1440
QY 1441 ATTCTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
DB 1441 ATTCTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
QY 1501 GCATTTCAACAAACAGGACCATAGATTTATGAAACTGGAAGGTAAACTGTTAACTCTGG 1560
DB 1501 GCATTTCAACAAACAGGACCATAGATTTATGAAACTGGAAGGTAAACTGTTAACTCTGG 1560
QY 1561 ATTAACGGCGATAAAGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAATTCGAGAAGAT 1620
DB 1561 ATTAACGGCGATAAAGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGATCCGATAAGAAATTCGAGAAGAT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGATCCGATAAGAAATTCGAGAAGAT 1680
QY 1681 GCGGCACTGGCGATGGCTCTGACATTAATCTTTCTGGCACACGACCGCTTGGTGGCTAC 1740
DB 1681 GCGGCACTGGCGATGGCTCTGACATTAATCTTTCTGGCACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAATCAACCGCAAGCTGATTGCTTACCCTGCTGCTTT 1800
DB 1741 GCTCAATCTGGCTGTGGCTGAATCAACCGCAAGCTGATTGCTTACCCTGCTGCTTT 1800
QY 1801 CGTTTTACTGGGATCGGATGATTAACCGCAAGCTGATTGCTTACCCTGCTGCTTT 1860
DB 1801 CGTTTTACTGGGATCGGATGATTAACCGCAAGCTGATTGCTTACCCTGCTGCTTT 1860
QY 1861 GAAGCGTTATCGCTCATTTATAACAAAGATCTGTGCGCAACCGCCCAAAACCTGGAA 1920
DB 1861 GAAGCGTTATCGCTCATTTATAACAAAGATCTGTGCGCAACCGCCCAAAACCTGGAA 1920
QY 1921 GAGATCCCGGCTGGATTAAGAACTGAAAGCAAGTAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGGCTGGATTAAGAACTGAAAGCAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGAGCTGGCGGTGATACGCTGGCGGAAAGCG 2100
DB 2041 TATGAAAACGGCAAGTACGACATTAAGAGCTGGCGGTGATACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAAGCAACACATGAATGCAGACCGATTAC 2160
DB 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAAGCAACACATGAATGCAGACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCAACAGCGATGACCATCAACGCGCGGTGG 2220
DB 2161 TCCATCGCAGAGCTGCTTTTAAAGGCAACAGCGATGACCATCAACGCGCGGTGG 2220
QY 2221 GCATGGTCCAACTCGACACCGCAAGTGAATTTAGTGTAAACGGTACTGCCGACCTTC 2280
DB 2221 GCATGGTCCAACTCGACACCGCAAGTGAATTTAGTGTAAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGTCACCATCCAAACCGTTCTGCTGGCTGAGCGCAGGTATTAAACGCGCCAGT 2340

Db	2281		AAGGGTCAACCATCAAAACCGTTCGTTGGCGTGCTGAGCGCAGGATTAACGCGCGCCAGT	2340
Qy	2341		CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAACATATCTGCTGACTGATGAAGGTCG	2400
Db	2341		CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAACATATCTGCTGACTGATGAAGGTCG	2400
Qy	2401		GAAGCGGTTAATAAGACAACACCGCTGGGTGCGGTAGCGCTTGAAGTCTTACGAGGAAGAG	2460
Db	2401		GAAGCGGTTAATAAGACAACACCGCTGGGTGCGGTAGCGCTTGAAGTCTTACGAGGAAGAG	2460
Qy	2461		TTGCGAAAGATCCACGCTATTGCGGCCACCATGGAAAAACGCCAAGAGGTGAAATCATG	2520
Db	2461		TTGCGAAAGATCCACGCTATTGCGGCCACCATGGAAAAACGCCAAGAGGTGAAATCATG	2520
Qy	2521		CCGAACATCCCGCAGATGTCGCTTTCTGTGATGCCGTGCGTACTCGGTGATCAACGCC	2580
Db	2521		CCGAACATCCCGCAGATGTCGCTTTCTGTGATGCCGTGCGTACTCGGTGATCAACGCC	2580
Qy	2581		GCCAGCGGTGTCAGACTGTGTCGATGAAGCCCTGAAAGACGCCGACAGCTAAATTGAGGTCG	2640
Db	2581		GCCAGCGGTGTCAGACTGTGTCGATGAAGCCCTGAAAGACGCCGACAGCTAAATTGAGGTCG	2640
Qy	2641		AACAACAACAATTAACAATAACAACCTCCGGATCGAGGGAAAGGATTTTCAGAATTC	2700
Db	2641		AACAACAACAATTAACAATAACAACCTCCGGATCGAGGGAAAGGATTTTCAGAATTC	2700

Search completed: August 3, 2005, 04:22:00
Job time : 1317.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacaccatcgaaatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	2700	81.8	7475	2	US-08-971-036-1	
2	2700	81.8	7475	3	US-09-096-570-1	
3	2700	81.8	7475	3	US-09-265-617B-1	
c	4	1417.6	43.0	5926	3	US-09-027-169-3
	5	1362	41.3	5201	4	US-09-640-882-2
	6	1362	41.3	5201	4	US-09-640-882-3
	7	1204	36.5	4557	4	US-08-778-717-5
8	1200.8	36.4	3832	1	US-08-148-675A-2	
9	1199.2	36.3	5248	3	US-08-487-283A-18	
c	10	1199.2	36.3	5248	5	PCT-US96-05611A-21
c	11	1199.2	36.3	5312	4	US-10-263-103-35
c	12	1199.2	36.3	5443	2	US-08-929-967-1
c	13	1199.2	36.3	5502	4	US-09-702-705-785
c	14	1199.2	36.3	5502	4	US-09-736-457-785
c	15	1199.2	36.3	5502	4	US-09-614-124B-785
c	16	1199.2	36.3	5502	4	US-09-671-325-785
c	17	1199.2	36.3	5502	4	US-09-589-184-785
c	18	1199.2	36.3	5502	4	US-09-658-824-785
c	19	1199.2	36.3	5502	2	US-08-929-967-3
c	20	1199.2	36.3	5873	4	US-09-695-437A-62
c	21	1199.2	36.3	6353	4	US-09-702-705-784
c	22	1199.2	36.3	6353	4	US-09-736-457-784
c	23	1199.2	36.3	6353	4	US-09-614-124B-784
c	24	1199.2	36.3	6353	4	US-09-671-325-784
c	25	1199.2	36.3	6353	4	US-09-589-184-784
c	26	1199.2	36.3	6353	4	US-09-658-824-784
27	1199.2	36.3	6361	3	US-08-646-538-7	
					Sequence 1, Appli	
					Sequence 1, Appli	
					Sequence 2, Appli	
					Sequence 3, Appli	
					Sequence 3, Appli	
					Sequence 4, Appli	
					Sequence 4, Appli	
					Sequence 5, Appli	
					Sequence 5, Appli	
					Sequence 6, Appli	
					Sequence 6, Appli	
					Sequence 7, Appli	
					Sequence 7, Appli	
					Sequence 8, Appli	
					Sequence 8, Appli	
					Sequence 9, Appli	
					Sequence 9, Appli	
					Sequence 10, Appli	
					Sequence 10, Appli	
					Sequence 11, Appli	
					Sequence 11, Appli	
					Sequence 12, Appli	
					Sequence 12, Appli	
					Sequence 13, Appli	
					Sequence 13, Appli	
					Sequence 14, Appli	
					Sequence 14, Appli	
					Sequence 15, Appli	
					Sequence 15, Appli	
					Sequence 16, Appli	
					Sequence 16, Appli	
					Sequence 17, Appli	
					Sequence 17, Appli	
					Sequence 18, Appli	
					Sequence 18, Appli	
					Sequence 19, Appli	
					Sequence 19, Appli	
					Sequence 20, Appli	
					Sequence 20, Appli	
					Sequence 21, Appli	
					Sequence 21, Appli	
					Sequence 22, Appli	
					Sequence 22, Appli	
					Sequence 23, Appli	
					Sequence 23, Appli	
					Sequence 24, Appli	
					Sequence 24, Appli	
					Sequence 25, Appli	
					Sequence 25, Appli	
					Sequence 26, Appli	
					Sequence 26, Appli	
					Sequence 27, Appli	
					Sequence 27, Appli	

c	28	1199.2	36.3	6361	3	US-09-503-222-7
c	29	1199.2	36.3	6727	2	US-08-125-462-2
c	30	1199.2	36.3	6727	2	US-08-891-848-2
c	31	1199.2	36.3	6799	2	US-08-125-462-5
c	32	1199.2	36.3	6799	2	US-08-891-848-5
c	33	1199.2	36.3	7659	3	US-09-128-314-4
c	34	1199.2	36.3	7676	3	US-09-056-556-213
c	35	1199.2	36.3	7676	3	US-09-072-596-208
c	36	1199.2	36.3	7676	4	US-09-072-967-213
c	37	1199.2	36.3	7676	4	US-09-287-849-9
c	38	1199.2	36.3	8031	3	US-09-643-597-254
c	39	1199.2	36.3	8031	4	US-09-480-884A-254
c	40	1199.2	36.3	8031	4	US-09-542-615A-254
c	41	1199.2	36.3	8031	4	US-09-606-421B-254
c	42	1199.2	36.3	8031	4	US-09-476-496A-254
c	43	1199.2	36.3	8031	4	US-09-630-940B-254
c	44	1199.2	36.3	8157	3	US-09-128-314-3
c	45	1199.2	36.3	8501	3	US-08-793-900-1

ALIGNMENTS

RESULT 1
US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8%; Score 2700; DB 2; Length 7475;

Best Local Similarity 100.0%; Pred. No. 0;				Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CGACACCATCGAATGTCGCAAAACCTTTCGGGTATGCGATGATAGCCCGGAGAGA	60				
Db	1	CGACACCATCGAATGTCGCAAAACCTTTCGGGTATGCGATGATAGCCCGGAGAGA	60				
QY	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGCGAGATGCGG	120				
Db	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGCGAGATGCGG	120				
QY	121	GTGTCCTTATCAGACCGTTTCGGGTGGTGAACACAGCCAGCCAGTTTCTGGAAAA	180				
Db	121	GTGTCCTTATCAGACCGTTTCGGGTGGTGAACACAGCCAGCCAGTTTCTGGAAAA	180				
QY	181	CGCGGAAAAAGTGGAAAGCGGAGTGGCGGCTGAAATTAACATTCCTCCACCGCGTGGC	240				
Db	181	CGCGGAAAAAGTGGAAAGCGGAGTGGCGGCTGAAATTAACATTCCTCCACCGCGTGGC	240				
QY	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGGCTTGCCACCTCCAGTCTGGCCCTGC	300				
Db	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGGCTTGCCACCTCCAGTCTGGCCCTGC	300				
QY	301	ACGCGCGTCGCAAAATGTCGCGGCGATTAATCTCGCGCGGATCAACTGGGTGCGACG	360				
Db	301	ACGCGCGTCGCAAAATGTCGCGGCGATTAATCTCGCGCGGATCAACTGGGTGCGACG	360				
QY	361	TGGTGGTTCGATGTTAGAACGAGCGGCTGAAAGCTGTAAGCTTCCGATCGGATGCGA	420				
Db	361	TGGTGGTTCGATGTTAGAACGAGCGGCTGAAAGCTGTAAGCTTCCGATCGGATGCGA	420				
QY	421	TTCTCGCGAACCGCTCAGTGGGCTGATCAATTAATCTCCGATCGGATGACGAGTGCCA	480				
Db	421	TTCTCGCGAACCGCTCAGTGGGCTGATCAATTAATCTCCGATCGGATGACGAGTGCCA	480				
QY	481	TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGGTATTTCTTGATGTTCTCGACCGA	540				
Db	481	TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGGTATTTCTTGATGTTCTCGACCGA	540				
QY	541	CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGCGATCGGCGGTGAGGATC	600				
Db	541	CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGCGATCGGCGGTGAGGATC	600				
QY	601	TGTCGCAATTCGGTCACAGCAAAATCGCGCTGTTAGCGGCGCAATTAAGTTCTCTCGG	660				
Db	601	TGTCGCAATTCGGTCACAGCAAAATCGCGCTGTTAGCGGCGCAATTAAGTTCTCTCGG	660				
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720				
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720				
QY	721	CGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780				
Db	721	CGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780				
QY	781	ATGAGGCGATGTTCCCACTCGCATGCTGTTAGCGGCGCAATTAAGTTCTCTCGG	840				
Db	781	ATGAGGCGATGTTCCCACTCGCATGCTGTTAGCGGCGCAATTAAGTTCTCTCGG	840				
QY	841	TGCGCGCATTTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG	900				
Db	841	TGCGCGCATTTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG	900				
QY	901	ACGATACGAGACAGTCTATGTTATATCCGCGGTTAACCATCAACAGGATTTTC	960				
Db	901	ACGATACGAGACAGTCTATGTTATATCCGCGGTTAACCATCAACAGGATTTTC	960				
QY	961	GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020				
Db	961	GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020				
QY	1021	AGGCAATACGCTGTTGCCGCTCTACTGTGTAAGAAAAACCACTTGGCGCCCAATA	1080				
Db	1021	AGGCAATACGCTGTTGCCGCTCTACTGTGTAAGAAAAACCACTTGGCGCCCAATA	1080				

Db	1021	AGGCAATACGCTGTTGCCGCTCTACTGTGTAAGAAAAACCACTTGGCGCCCAATA	1080				
QY	1081	CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140				
Db	1081	CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140				
QY	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTATTAG	1200				
Db	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTATTAG	1200				
QY	1201	GCACAAATCTCATCTTTGACAGCTTATCATCGACTGACCGTGCACCAATGCTTCTGGG	1260				
Db	1201	GCACAAATCTCATCTTTGACAGCTTATCATCGACTGACCGTGCACCAATGCTTCTGGG	1260				
QY	1261	TCAGGAGCCATCGAAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCAATAATTCG	1320				
Db	1261	TCAGGAGCCATCGAAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCAATAATTCG	1320				
QY	1321	TGTCGCTCAAGGCGCACTCCGCTTCTGGATAATGTTTTTTTGGCGCGACATCAACCGTT	1380				
Db	1321	TGTCGCTCAAGGCGCACTCCGCTTCTGGATAATGTTTTTTTGGCGCGACATCAACCGTT	1380				
QY	1381	CTGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGGA	1440				
Db	1381	CTGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGGA	1440				
QY	1441	ATTGTAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500				
Db	1441	ATTGTAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500				
QY	1501	GCATTCACCAACAAAGGACCATAGATTATGAAACTGAAAGAGGTAACTGGTAACTGCG	1560				
Db	1501	GCATTCACCAACAAAGGACCATAGATTATGAAACTGAAAGAGGTAACTGGTAACTGCG	1560				
QY	1561	ATTAAACGGCGATAAAGCTATAACGGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT	1620				
Db	1561	ATTAAACGGCGATAAAGCTATAACGGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT	1620				
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680				
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680				
QY	1681	CGCGCAACTGGCGATCGCCCTGACATTAATCTCTGGGCAACAGCCGCTTGGTGCTAC	1740				
Db	1681	CGCGCAACTGGCGATCGCCCTGACATTAATCTCTGGGCAACAGCCGCTTGGTGCTAC	1740				
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAGCTGTAT	1800				
Db	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAGCTGTAT	1800				
QY	1801	CCGTTTACCTGGGATCGCGTACGTTACACGGCAAGCTGATTTGCTTACCGATCGCTGTT	1860				
Db	1801	CCGTTTACCTGGGATCGCGTACGTTACACGGCAAGCTGATTTGCTTACCGATCGCTGTT	1860				
QY	1861	GAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCAAAAACCTGGGAA	1920				
Db	1861	GAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCAAAAACCTGGGAA	1920				
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAGAGGAAAGTGAAGAGCGCTGTATGTTTCAAC	1980				
Db	1921	GAGATCCCGCGCTGGATAAGAACTGAAGAGGAAAGTGAAGAGCGCTGTATGTTTCAAC	1980				
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGCTGACGGGGTTTATCGGTTCAAG	2040				
Db	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGCTGACGGGGTTTATCGGTTCAAG	2040				
QY	2041	TATGAACCGCAAGTACGACATTAAGAGCTGGCGTGGATAACCGTGGCGGAAAGCG	2100				
Db	2041	TATGAACCGCAAGTACGACATTAAGAGCTGGCGTGGATAACCGTGGCGGAAAGCG	2100				
QY	2101	GGTCTGACCTTCTGTTGACCTGATTAACCAACACATGAATGACAGACCCGATTAC	2160				
Db	2101	GGTCTGACCTTCTGTTGACCTGATTAACCAACACATGAATGACAGACCCGATTAC	2160				

QY 2161 TCCATCGCAGAGCTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTATGGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTATGGTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCAATCAACACCGTTGCTGGCGTGTGAGCGCAGGATTAACGCGCCGCTG 2340
Db 2281 AAGGTCACCAATCAACACCGTTGCTGGCGTGTGAGCGCAGGATTAACGCGCCGCTG 2340
QY 2341 CGAACAAGAGCTGCGAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGCTCTG 2400
Db 2341 CGAACAAGAGCTGCGAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGCTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGCTATTGCCGCCACCATGGAAACCGCCAGAAAGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATTGCCGCCACCATGGAAACCGCCAGAAAGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGACACTGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGACACTGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTTCAGAAATTC 2700

RESULT 2
US-09-096-570-1
; Sequence 1, Application US/09096570
; Patent No. 6018020
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/096,570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522

; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-09-096-570-1

Query Match 81.8%; Score 2700; DB 3; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGTAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGTAGCGCCGGAAGAGA 60
QY 61 GTC AATTCAGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTCGAGAGTATGCGG 120
Db 61 GTC AATTCAGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTCGAGAGTATGCGG 120
QY 121 GTG TCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCAGCCGTTTCTGCGAAAA 180
Db 121 GTG TCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCAGCCGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGCTGGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGCTGGGCAC 240
QY 241 AACAACTCGCGCGGCAACAGTCGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTCGCGCGGCAACAGTCGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCCGTCGCAAAATTTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCCGTCGCAAAATTTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGATAGAAAGCGCGCTGCAAGCTCTGAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTGTGATGATAGAAAGCGCGCTGCAAGCTCTGAAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATTAATCTATCCGCTGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATTAATCTATCCGCTGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGGGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGGGACTGGGCGTGGAGCATC 600
QY 601 TGGTGGCATTTGGTCCACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGGCATTTGGTCCACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGGACGTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGGACGTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGCATGATGCGGCTGGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGCATGATGCGGCTGGGGCGCAA 840

841 TGCGGCCATTACCGAGTCCGGCTCGCGTGGTGGTGGGATATCTCGGTAGTGGGATAG 900
841 TGCGCGCATACCGAGTCCGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATAG 900
901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAACCCACCTCAAAACAGGATTTTC 960
901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAACCCACCTCAAAACAGGATTTTC 960
961 GCTGCTGGGGCAACACGCGTGGACCGCTTCTGCACTCTCTCAGGCGCCAGGCGGTGA 1020
961 GCTGCTGGGGCAACACGCGTGGACCGCTTCTGCACTCTCTCAGGCGCCAGGCGGTGA 1020
1021 AGGCAATCAGCTGTTCGCCGCTCATCTGTTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
1021 AGGCAATCAGCTGTTCGCCGCTCATCTGTTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGCTGCGACGACAGGTTT 1140
1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGCTGCGACGACAGGTTT 1140
1141 CCCGACTGAAAGCGGAGTGAAGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
1141 CCCGACTGAAAGCGGAGTGAAGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
1201 GCACAATTCATGTTTGAAGCTTATCATGACTGCAAGTGCACCAATGCTTCTGGCG 1260
1201 GCACAATTCATGTTTGAAGCTTATCATGACTGCAAGTGCACCAATGCTTCTGGCG 1260
1261 TCAGGACGCAATCGAAGCTGTTGATGCTGTCGAGTGCACCAATGCTTCTGGCG 1320
1261 TCAGGACGCAATCGAAGCTGTTGATGCTGTCGAGTGCACCAATGCTTCTGGCG 1320
1321 TGTGCTCAAGCGCACTCCCGTTCGATTAATGTTTTCGCGGACATCATACCGTT 1380
1321 TGTGCTCAAGCGCACTCCCGTTCGATTAATGTTTTCGCGGACATCATACCGTT 1380
1381 CTGGCAATATCTGAAATGCTGTTGCAATTAATCATCGCTCGTATATGTTGGA 1440
1381 CTGGCAATATCTGAAATGCTGTTGCAATTAATCATCGCTCGTATATGTTGGA 1440
1441 ATTGTAGCGGATTAACAAATTTACAGAGAACAGCAGTCCGTTAGTGGTTTTCAGGA 1500
1441 ATTGTAGCGGATTAACAAATTTACAGAGAACAGCAGTCCGTTAGTGGTTTTCAGGA 1500
1501 GCATTTACCAACAGGACCATAGATTATGAAATCTGAAGAGTAACTGGTAACTCGG 1560
1501 GCATTTACCAACAGGACCATAGATTATGAAATCTGAAGAGTAACTGGTAACTCGG 1560
1561 ATTAACGCGGATAAAGGCTAATACGCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
1561 ATTAACGCGGATAAAGGCTAATACGCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
1681 CGCGCACTGGCGATGCGCTGACATTAATCTTGGGCAACAGCAGTGGTGGCTAC 1740
1681 CGCGCACTGGCGATGCGCTGACATTAATCTTGGGCAACAGCAGTGGTGGCTAC 1740
1741 GCTCAATCTGGCTGTTGGTGAATCACCCTGGACAAAGCGTTCCAGGCAAGCTGTAT 1800
1741 GCTCAATCTGGCTGTTGGTGAATCACCCTGGACAAAGCGTTCCAGGCAAGCTGTAT 1800
1801 CCGTTTACCTGGATGCGTACGTTAACAACGCAAGCTGATGCTTACCGATCGCTGTT 1860
1801 CCGTTTACCTGGATGCGTACGTTAACAACGCAAGCTGATGCTTACCGATCGCTGTT 1860
1861 GAAGCGTTATCGTGAATTTATACAGAGTCTGCTGCCGAAACCCGCAAAACCTCGGAA 1920
1861 GAAGCGTTATCGTGAATTTATACAGAGTCTGCTGCCGAAACCCGCAAAACCTCGGAA 1920

1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAGAGCGGCTGATGTTCAAC 1980
1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAGAGCGGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTCCTGCTGACGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTCCTGCTGACGGGGTTATGCGTTCAAG 2040
2041 TATCAAAAACGCAAGTACGACATTAAGAGAGTGGGCGTGGATACGCTGGCGGCAAGCG 2100
2041 TATCAAAAACGCAAGTACGACATTAAGAGAGTGGGCGTGGATACGCTGGCGGCAAGCG 2100
2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACATGATGAGACACCGATTAC 2160
2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACATGATGAGACACCGATTAC 2160
2161 TCCATCCGAGAAGCTGCTTTTAAAGAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
2161 TCCATCCGAGAAGCTGCTTTTAAAGAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
2221 GCATGCTCAACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC 2280
2221 GCATGCTCAACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC 2280
2281 AAGGTCACCAACATCCCAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGGCCGAGT 2340
2281 AAGGTCACCAACATCCCAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGGCCGAGT 2340
2341 CCGAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
2341 CCGAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGCTGAGCGTGAAGTCTTACGAGGAAG 2460
2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGCTGAGCGTGAAGTCTTACGAGGAAG 2460
2461 TTGCGCAAGAGTCCAGTATTTGCGCCACCATGGAAGCGCCAGAAAGGTGAATCATG 2520
2461 TTGCGCAAGAGTCCAGTATTTGCGCCACCATGGAAGCGCCAGAAAGGTGAATCATG 2520
2521 CCGAACATCCGCGAGATGTCGCTTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
2521 CCGAACATCCGCGAGATGTCGCTTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
2581 GCAGCGGCTGCTGAGTCTGCTGAGTGAAGCGCTGGAAGACCGCGAGACTAATTCGAGCTCG 2640
2581 GCAGCGGCTGCTGAGTCTGCTGAGTGAAGCGCTGGAAGACCGCGAGACTAATTCGAGCTCG 2640
2641 AACAAACAAACAAATAAATAACAAACAAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
2641 AACAAACAAACAAATAAATAACAAACAAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700

RESULT 3
US-09-265-617B-1
; Sequence 1, Application US/09265617B
; Patent No. 6372883
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R.
; APPLICANT: Hurst, David N.
; APPLICANT: Jones, Philip S.
; APPLICANT: Kay, Paul B.
; APPLICANT: Raynham, Tony M.
; APPLICANT: Wilson, Francis X.
; TITLE OF INVENTION: Antiviral Medicaments
; FILE REFERENCE: 20052 antiviral medicaments
; CURRENT APPLICATION NUMBER: US/09/265,617B
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: GB9806815.8
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

; LENGTH: 7475

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Vector and

; OTHER INFORMATION: Gene Fragments

US-09-265-617B-1

Query Match

Best Local Similarity 81.8%; Score 2700; DB 3; Length 7475;

Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGGACACCATCGAATGGTGGCAAAACCTTTCGGGGTATGCGCATGATAGCGCCGGAAGAGA	60
Dd	1	CGGACACCATCGAATGGTGGCAAAACCTTTCGGGGTATGCGCATGATAGCGCCGGAAGAGA	60
Qy	61	GTCATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATCTCGCAGAGTATGCCG	120
Dd	61	GTCATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATCTCGCAGAGTATGCCG	120
Qy	121	GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180
Dd	121	GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180
Qy	181	CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTCGCCAACCGCGTGGCAC	240
Dd	181	CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTCGCCAACCGCGTGGCAC	240
Qy	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC	300
Dd	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC	300
Qy	301	ACGCGCGTTCGCAAAATGTTCGGCGGATTTAAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
Dd	301	ACGCGCGTTCGCAAAATGTTCGGCGGATTTAAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
Qy	361	TGGTGGTTCGATGTAGAACGAAAGCGGCTCGAAGCGCTGTAAGCGGCGGTGCAATATC	420
Dd	361	TGGTGGTTCGATGTAGAACGAAAGCGGCTCGAAGCGCTGTAAGCGGCGGTGCAATATC	420
Qy	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
Dd	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
Qy	481	TTGCTGTGGAAAGCTGCCTGCAATAATGTTCCGGCGTATTTCTTGATGTCTCTGACCAGA	540
Dd	481	TTGCTGTGGAAAGCTGCCTGCAATAATGTTCCGGCGTATTTCTTGATGTCTCTGACCAGA	540
Qy	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC	600
Dd	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC	600
Qy	601	TGGTGGCATTTGGTCCACGAAATCGCGCTGTAGCGGGCCCATTAAGTCTGCTCGG	660
Dd	601	TGGTGGCATTTGGTCCACGAAATCGCGCTGTAGCGGGCCCATTAAGTCTGCTCGG	660
Qy	661	CGCGTCTGCGTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Dd	661	CGCGTCTGCGTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAAACGGGAACGGCATGGAATGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA	780
Dd	721	CGGAAACGGGAACGGCATGGAATGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA	780
Qy	781	ATGAGGGCATCGTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCA	840
Dd	781	ATGAGGGCATCGTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCA	840
Qy	841	TGCGGCGCAATACCGAGTCCGGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATAG	900
Dd	841	TGCGGCGCAATACCGAGTCCGGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATAG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAAGGATTTTC	960

Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAAGGATTTTC	960
Qy	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
Dd	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTTGCCGCTCTCATCTGCTGAAAGAAACCAACCCCTGGCGGCCAATA	1080
Dd	1021	AGGGCAATCAGCTGTTGCCGCTCTCATCTGCTGAAAGAAACCAACCCCTGGCGGCCAATA	1080
Qy	1081	CGAAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCAGTGCACGACAGGTTT	1140
Dd	1081	CGAAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCAGTGCACGACAGGTTT	1140
Qy	1141	CCGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
Dd	1141	CCGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGCTGCACGGTGCACCAATGCTTCTGGCG	1260
Dd	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGCTGCACGGTGCACCAATGCTTCTGGCG	1260
Qy	1261	TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGGTCGTAATCACTGCATAATTCG	1320
Dd	1261	TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGGTCGTAATCACTGCATAATTCG	1320
Qy	1321	TGTGCTCAAGCGCGCACTCCCGTCTTGATTAATGTTTTTGGCGCGACATCAATACGGTT	1380
Dd	1321	TGTGCTCAAGCGCGCACTCCCGTCTTGATTAATGTTTTTGGCGCGACATCAATACGGTT	1380
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATATGTGTGA	1440
Dd	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATATGTGTGA	1440
Qy	1441	ATTGTGAGCGGATTAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA	1500
Dd	1441	ATTGTGAGCGGATTAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA	1500
Qy	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAAGAGTAAACTGGTAAATCTCGG	1560
Dd	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAAGAGTAAACTGGTAAATCTCGG	1560
Qy	1561	ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTGGTGAAGAAATTCGAGAAAGAT	1620
Dd	1561	ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTGGTGAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGGAATTAAGTCAACCGTTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
Dd	1621	ACCGGAATTAAGTCAACCGTTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
Qy	1681	GGCGCACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Dd	1681	GGCGCACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGCGCTGTTGGCTGAAAATCAACCGGACAAAGCGTTCAGGACAAAGCTGAT	1800
Dd	1741	GCTCAATCTGCGCTGTTGGCTGAAAATCAACCGGACAAAGCGTTCAGGACAAAGCTGAT	1800
Qy	1801	CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCGGATCGCTGTT	1860
Dd	1801	CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCGGATCGCTGTT	1860
Qy	1861	GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCCAAAACCTGGGAA	1920
Dd	1861	GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCCAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGAGTAAAGAGCGCGCTGATGTTCAAC	1980
Dd	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGAGTAAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGGAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATCGGTTCAAG	2040

1981	DB	CTGCAAGAACGTA	CTTCACTGGCCGCTG	ATGCTGCTGACGGGGGT	TATGCGTTCAAG	2040		
2041	QY	TATGAAAA	CGGCAAGTACG	ACATTAAGACGCTGGCGTGG	AATACGCTGGCGCGAAGACG	2100		
2041	DB	TATGAAAA	CGGCAAGTACG	ACATTAAGACGCTGGCGTGG	ATACGCTGGCGCGAAGACG	2100		
2101	QY	GGTCTGACCTT	CCCTGGTTGAC	CTGATTA	AAAAAACAACATGAATG	CAGACACCGGATTAC	2160	
2101	DB	GGTCTGACCTT	CCCTGGTTGAC	CTGATTA	AAAAAACAACATGAATG	CAGACACCGGATTAC	2160	
2161	QY	TCCATCGAGAAG	CTGCTTTAA	TAAAGGCGAAAC	ACGATGACCATCAACGCGCCG	GTGG	2220	
2161	DB	TCCATCGAGAAG	CTGCTTTAA	TAAAGGCGAAAC	ACGATGACCATCAACGCGCCG	GTGG	2220	
2221	QY	GCATGGTCCA	ACATCGAC	CACAGCAAAAGTGA	TTATGGTGAACGGTACTGCCGAC	CTTC	2280	
2221	DB	GCATGGTCCA	ACATCGAC	CACAGCAAAAGTGA	TTATGGTGAACGGTACTGCCGAC	CTTC	2280	
2281	QY	AAGGGTCA	ACCATCAAA	CCGTTGTTGGCGT	GTGAGCGCAGGTATTAACGCGCG	CAGT	2340	
2281	DB	AAGGGTCA	ACCATCAAA	CCGTTGTTGGCGT	GTGAGCGCAGGTATTAACGCGCG	CAGT	2340	
2341	QY	CCGAA	CAAGAGCTGG	CAAAAGAGTTCTCG	ATAAACTATCTGCTGATGAAG	GTCTG	2400	
2341	DB	CCGAA	CAAGAGCTGG	CAAAAGAGTTCTCG	ATAAACTATCTGCTGATGAAG	GTCTG	2400	
2401	QY	GAAGCGGT	TAATAAAG	CAAAACCGCTGG	GTGCGCTGAAGCTTACGAGGA	AGAG	2460	
2401	DB	GAAGCGGT	TAATAAAG	CAAAACCGCTGG	GTGCGCTGAAGCTTACGAGGA	AGAG	2460	
2461	QY	TTGGCGAA	AGATCCA	CGTATTGCGGCA	CCATGGAAAA	CGCCAGAAAGGTGAAATCATG	2520	
2461	DB	TTGGCGAA	AGATCCA	CGTATTGCGGCA	CCATGGAAAA	CGCCAGAAAGGTGAAATCATG	2520	
2521	QY	CCGAA	CATCCCGAG	ATGTCGGCTTTCTG	TATGCCGTGCGTACTCGGTG	ATCAACGCC	2580	
2521	DB	CCGAA	CATCCCGAG	ATGTCGGCTTTCTG	TATGCCGTGCGTACTCGGTG	ATCAACGCC	2580	
2581	QY	GCCAGCGT	CGTGCAG	ACTGTGATGAAG	CCCTGAAAGA	CGCGCAGACTAATTCGAGCTCG	2640	
2581	DB	GCCAGCGT	CGTGCAG	ACTGTGATGAAG	CCCTGAAAGA	CGCGCAGACTAATTCGAGCTCG	2640	
2641	QY	AACA	CAACAACAA	TAAACAACTCG	GGATCAAC	ACCTCGGGATCGAGGGAAGATTTTCAGAA	TTTC	2700
2641	DB	AACA	CAACAACAA	TAAACAACTCG	GGATCAAC	ACCTCGGGATCGAGGGAAGATTTTCAGAA	TTTC	2700

RESULT 4

```

US-09-027-169-3/c
/ Sequence 3, Application US/09027169
/ Patent No. 6420524
/ GENERAL INFORMATION:
/ APPLICANT: CRAIG, NANCY L
/ TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
/ TITLE OF INVENTION: ATP-DEPENDENT TRANSCRIPTION PROTEINS
/ NUMBER OF SEQUENCES: 15
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Anne Brown (Alston & Bird, LLP)
/ STREET: 3605 Glenwood Ave.
/ CITY: Raleigh
/ STATE: NC
/ COUNTRY: USA
/ ZIP: 27608
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/027,169
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:

```

|||||
Db 784 GAGGGCATGTTCCCACTGCGATGCTGGTTGGTCCAAACGATCAGATGGCGTGGCGCAATG 725
Qy CGCGCATTTACCGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGATACGAC 902
Db CGCGCATTTACCGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGATACGAC 665
Qy GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGAGATTTTCG 962
Db GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGAGATTTTCG 605
Qy CTGCTGGGGCAACAGAGTGGACCGTTGCTGCAACTCTCTCAGGGCCAGSCGGTGAAG 1022
Db CTGCTGGGGCAACAGAGTGGACCGTTGCTGCAACTCTCTCAGGGCCAGSCGGTGAAG 545
Qy GCGAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCACTGGCGCCCAATACG 1082
Db GCGAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCACTGGCGCCCAATACG 485
Qy CAAACCGCTCTCCCGCGCGTTGGCGGATTCATTAATGACAGCTGGCAGACAGGTTTCC 1142
Db CAAACCGCTCTCCCGCGCGTTGGCGGATTCATTAATGACAGCTGGCAGACAGGTTTCC 425
Qy CGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGC-----GC 1202
Db CGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGC-----GC 376
Qy ACAATTTCTCATGTTTGACAGCTTATCATCGACTGCACCGTGACCAATGCTTCTGGCGTC 1262
Db GAAATTGATCTGGTTTGACAGCTTATCATCGACTGCACCGTGACCAATGCTTCTGGCGTC 316
Qy AGGCAACCATCGGAAGCTGTGGTATGGCTGTCGAGTGCATTAATCACTGCATAATTCGTG 1322
Db AGGCAACCATCGGAAGCTGTGGTATGGCTGTCGAGTGCATTAATCACTGCATAATTCGTG 256
Qy TCGCTCAAGGCGCACTCCGTTCTCGATAATGTTTTTGGCGCGACATCAACCGTTCT 1382
Db TCGCTCAAGGCGCACTCCGTTCTCGATAATGTTTTTGGCGCGACATCAACCGTTCT 196
Qy GCGAAATATTTCTGAATGAGCTGTTGACAAATTAATCAT- CGGCTCGTATAATGTTGGAA 1441
Db GCGAAATATTTCTGAATGAGCTGTTGACAAATTAATCATCAATCAATCGGCTCGTATAATGTTGGAA 136
Qy TTGTGAGCGGATAACAATTTTCAACAGGAAACAGCC 1477
Db TTGTGAGCGGATAACAATTTTCAACAGGAAACAGAC 100

RESULT 5

US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE2
US-09-640-882-2

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
Qy 3 GACACCATCGAATGTCGCAAAACCTTTTCGGGTATGGCATATAGTCGCGCGCAAGAGAGT 62
Db GACACCATCGAATGTCGCAAAACCTTTTCGGGTATGGCATATAGTCGCGCGCGCAAGAGAGT 925
Qy 63 CAATTCAGGGTGGTAATGTGAAACCACTAAACGTTATACGATGTGCGAGAGTATGCCGGT 122
Db CAATTCAGGGTGGTAATGTGAAACCACTAAACGTTATACGATGTGCGAGAGTATGCCGGT 985
Qy 123 GTCTCTTATCAGACCGTTTCCGGGTGCTGAAACGAGCCAGCCAGCTTCTTCGGAACAG 182
Db GTCTCTTATCAGACCGTTTCCGGGTGCTGAAACGAGCCAGCCAGCTTCTTCGGAACAG 1045
Qy 183 CGGGAAGGTTGGAAGCGCGATGGCGGAGCTGAATTAATATATTCCTCCAAACCGGTGCAAA 242
Db CGGGAAGGTTGGAAGCGCGATGGCGGAGCTGAATTAATATATTCCTCCAAACCGGTGCAAA 1105
Qy 243 CAATTCAGGGTGGTAATGTGAAACCACTAAACGTTATACGATGTGCGAGTCTGGCCCTGCAC 302
Db CAATTCAGGGTGGTAATGTGAAACCACTAAACGTTATACGATGTGCGAGTCTGGCCCTGCAC 1165
Qy 303 GCGCGTGCCTAAATGTCGCGCGATTAATATCTCGCGCGATCAACTGGGTGCCAGCGTG 362
Db GCGCGTGCCTAAATGTCGCGCGATTAATATCTCGCGCGATCAACTGGGTGCCAGCGTG 1225
Qy 363 GTGGTGTGATGTTAGAAAGCGGCGTCAAGGCTGTAAAGCGGCGGTGCAATCTT 422
Db GTGGTGTGATGTTAGAAAGCGGCGTCAAGGCTGTAAAGCGGCGGTGCAATCTT 1285
Qy 423 CTCGCGCAACCGTCACTGGGCTGATCAATTAATCTATCCGTGGATGACAGGATGCCATT 482
Db CTCGCGCAACCGTCACTGGGCTGATCAATTAATCTATCCGTGGATGACAGGATGCCATT 1345
Qy 483 GCTGTGGAAGTGCCTGCATTAATGTTCCGGGTATTTCTTGATGTCTCTGACAGACA 542
Db GCTGTGGAAGTGCCTGCATTAATGTTCCGGGTATTTCTTGATGTCTCTGACAGACA 1405
Qy 543 CCCATCAACAGTATTATTTCTCCCATGAAGCGTACGCGACTGGCGGTGGAGCATCTG 602
Db CCCATCAACAGTATTATTTCTCCCATGAAGCGTACGCGACTGGCGGTGGAGCATCTG 1465
Qy 603 CTCGCAATTTGGTCAACAGCAATCGCGTGTGTAGCGGCGCCATTAAGTTCTGTCTCGCG 662
Db CTCGCAATTTGGTCAACAGCAATCGCGTGTGTAGCGGCGCCATTAAGTTCTGTCTCGCG 1525
Qy 663 CGTCTGCGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db CGTCTGCGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAAAGGGAAGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGAAT 782
Db GAAAGGGAAGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGAAT 1645
Qy 783 GAGGCGATCGTTCCCATCTGCGATGCTGTTGTCACACGATCAGATGCGCTGGCGCAATG 842
Db GAGGCGATCGTTCCCATCTGCGATGCTGTTGTCACACGATCAGATGCGCTGGCGCAATG 1705
Qy 843 CGCGCATTTACGAGTCCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACGAC 902
Db CGCGCATTTACGAGTCCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGCAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTCG 962
Db GATACCGAAGCAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTCG 1825
Qy 963 CTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1022
Db CTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1885
Qy 1023 GCGAATCAGTGTGTCGCGCTCTCACTGTGTAAGAAACCAACCCCTGGCGCCCAATACG 1082

|||||
Db 2184 TCGCTCAAGGCCACTCCGTTCTGATATGTTTTTTCGCCGACATCATAAACGGTTCT 2243
QY 1383 GGCAAAATATCT 1394
|||||
Db 2244 GGCAAAATATCT 2255
|||||
RESULT 7
US-08-778-717-5
; Sequence 5, Application US/08778717
; Patent No. 6602689
; GENERAL INFORMATION:
; APPLICANT: UENO, EIICHI
; APPLICANT: NOBUYUKI, FUJII
; APPLICANT: OKADA, MASAHISA
; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN
; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR
; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESS: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/778,717
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 352225/1995
; FILING DATE: 28-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2084-031-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4557 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; ORIGINAL SOURCE:
; ORGANISM: E. COLI
; STRAIN: BL21 (DE3)
; PUBLICATION INFORMATION:
; AUTHORS: NOBUYUKI FUJII ET AL,
; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED
; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF
; TITLE: EXPRESSING SAID FUSED PROTEIN
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557
US-08-778-717-5
Query Match 36.5%; Score 1204; DB 4; Length 4557;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACCATCAATGGTCAAAACCTTTCGGGGTATGCATGATAGCGCCGGAAGAGA 60
Db 2563 CCGACACCATCAATGGTCAAAACCTTTCGGGGTATGCATGATAGCGCCGGAAGAGA 2622

QY 61 GTCAATTCCAGGTCGTGAATGTAACCAAGTAAAGTTATACGATGTCGACAGTATCGG 120
Db 2623 GTCAATTCCAGGTCGTGAATGTAACCAAGTAAAGTTATACGATGTCGACAGTATCGG 2682
QY 121 GTGTCCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
Db 2683 GTGTCCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 2742
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGTGAATTAATTCCTCCAAACCGCTGGCAC 240
Db 2743 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGTGAATTAATTCCTCCAAACCGCTGGCAC 2802
QY 241 AACAACTGGCGGGGAAAAACAGTCGTTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 2803 AACAACTGGCGGGGAAAAACAGTCGTTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 2862
QY 301 ACGGCGCTCGCAAAATTCGCGGCGATTAATCTCGCGCGGATCAATCACTGGGTGCGACGG 360
Db 2863 ACGGCGCTCGCAAAATTCGCGGCGGATTAATCTCGCGCGGATCAATCACTGGGTGCGACGG 2922
QY 361 TGGTGGTCTCGATGGTAGAACGAGCGCGTGAAGCCCTGAAAGCGCGGTGCAACAATC 420
Db 2923 TGGTGGTCTCGATGGTAGAACGAGCGCGTGAAGCCCTGAAAGCGCGGTGCAACAATC 2982
QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATTCGCTGGATGACACGAGTACCA 480
Db 2983 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATTCGCTGGATGACACGAGTACCA 3042
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGCGCTTATTTCTTGATGTCCTGACCCAGA 540
Db 3043 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGCGCTTATTTCTTGATGTCCTGACCCAGA 3102
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 3103 CACCCATCAACAGTATTAATTTCTCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 3162
QY 601 TGGTCGCAATTCGGTCACCAAGCAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db 3163 TGGTCGCAATTCGGTCACCAAGCAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG 3222
QY 661 CGCGTCTCGCTGCGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 3223 CGCGTCTCGCTGCGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3282
QY 721 CGAAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 3283 CGAAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 3342
QY 781 ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGGCA 840
Db 3343 ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGGCA 3402
QY 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATAG 900
Db 3403 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATAG 3462
QY 901 ACGATACCGGAAGACAGCTCATGTTATATCCGCGCTTAAACCAACCATCAAAAGGATTTTC 960
Db 3463 ACGATACCGGAAGACAGCTCATGTTATATCCGCGCTTAAACCAACCATCAAAAGGATTTTC 3522
QY 961 GCCTCTGGGGCAACACGAGTGGACCGCTTGTGCGAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 3523 GCCTCTGGGGCAACACGAGTGGACCGCTTGTGCGAACTCTCTCAGGCGCAGGCGGTGA 3582
QY 1021 AGGCAATCAGTGTGCGCGCTCTCACTGTTGMAAGAAACCAACCTGGCGCCCAATA 1080
Db 3583 AGGCAATCAGTGTGCGCGCTCTCACTGTTGMAAGAAACCAACCTGGCGCCCAATA 3642
QY 1081 CGCAACCGCGCTCTCCCGCGCGTTGGCGGATTCATTAATGAGTGGCAGCGACAGGTTT 1140
Db 3643 CGCAACCGCGCTCTCCCGCGCGTTGGCGGATTCATTAATGAGTGGCAGCGACAGGTTT 3702
QY 1141 CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTGTAGTCACTCACTTAG 1200

Db 3703 CCGAGTGGAAAGCGGCGAGTGGCGCAACGAATTAATGTAGTACTCATTAG 3762
Qy 1201 GCAC 1204
Db 3763 GCAC 3766

RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerra, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Teai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3832 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-148-675A-2

Query Match 36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAGAGA 60
Db 3 CCGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAGAGA 62

Qy 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTTATACGATGTCGCAGAGTATGCCG 120
Db 63 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTTATACGATGTCGCAGAGTATGCCG 122

Qy 121 GTGTCCTCTATCAGACCGTTTCCCGCGTGTGAACACGAGTGTGAACACGAGTGTGCGG 180
Db 123 GTGTCCTCTATCAGACCGTTTCCCGCGTGTGAACACGAGTGTGAACACGAGTGTGCGG 182

Qy 181 CGCGGGAAAAAGTGAAGCGCGATGGCGAGTGAATTACATTCCCAACCGCGTGGCAC 240
Db 183 CGCGGGAAAAAGTGAAGCGCGATGGCGAGTGAATTACATTCCCAACCGCGTGGCAC 242

Qy 241 AACNACTGGCGGGCAACAGTCGTTGCTGATTTGGCGTGGCCACTCCAGTCTGGCCCTGC 300

Db 243 AACAACTGGCGGCAACAGATCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 302
Qy 301 ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 303 ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 362
Qy 361 TGGTGGTGTGATGGTAGAAGCGGCTGTAAGCGCTGTAAGCGCGCGGTGCACAATC 420
Db 363 TGGTGGTGTGATGGTAGAAGCGGCTGTAAGCGCTGTAAGCGCGCGGTGCACAATC 422
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATCACTATCCGCTGGATGACACAGATGCCA 480
Db 423 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATCACTATCCGCTGGATGACACAGATGCCA 482
Qy 481 TTGCTGTGGAAGCTGCTGCACCTAAATGTCGGGGTTATTTCTTGATGTTCTTGACACAGA 540
Db 483 TTGCTGTGGAAGCTGCTGCACCTAAATGTCGGGGTTATTTCTTGATGTTCTTGACACAGA 542
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 543 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 602
Qy 601 TGGTGCATTGGGTCAACAGCAAAATCGCGTGTAGCGGCGCCATTAAAGTTCTGTCTCGG 660
Db 603 TGGTGCATTGGGTCAACAGCAAAATCGCGTGTAGCGGCGCCATTAAAGTTCTGTCTCGG 662
Qy 661 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 663 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 722
Qy 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 723 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGGTTTCAACAAACCATGCAAAATGCTGA 782
Qy 781 ATGAGGCGATCGTTCCCATGTCGATGCTGTTGCCAACCATGAGTGGCGTGGCGCAA 840
Db 783 ATGAGGCGATCGTTCCCATGTCGATGCTGTTGCCAACCATGAGTGGCGTGGCGCAA 842
Qy 841 TGGCGGCCATTACCGAGTCCGGCTCGGCTGGGTGGGATATCTCGGTAGTGGGATACG 900
Db 843 TGGCGGCCATTACCGAGTCCGGCTCGGCTGGGTGGGATATCTCGGTAGTGGGATACG 902
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGATTTTC 960
Db 903 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGATTTTC 962
Qy 961 GCCTGCTGGGGCAACACCGGTGGACCGCTTGTGCAACTCTCTAGGGCGAGCGGTGA 1020
Db 963 GCCTGCTGGGGCAACACCGGTGGACCGCTTGTGCAACTCTCTAGGGCGAGCGGTGA 1022
Qy 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
Db 1023 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1082
Qy 1081 CGCAAAACCGCTCTCCCCCGCGTTGGCGGATTCATTATGAGTGGCACGACAGGTTT 1140
Db 1083 CGCAAAACCGCTCTCCCCCGCGTTGGCGGATTCATTATGAGTGGCACGACAGGTTT 1142
Qy 1141 CCGCTGTGAAGCGGCGAGTGGCGGCAACGCAATTAATGAGTGTAGTCTCACTATTAG 1200
Db 1143 CCGCTGTGAAGCGGCGAGTGGCGGCAACGCAATTAATGAGTGTAGTCTCACTATTAG 1202
Qy 1201 GCAC 1204
Db 1203 GCAC 1206

RESULT 9
US-08-487-283A-18/c
; Sequence 18, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.

APPLICANT: Matis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Squinto, Stephen P.
APPLICANT: Thomas, Thomas C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: Seth A. Fidel
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Cetrus 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790
TELEFAX: (203) 772-3655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: pBT Trc SOS/NI
DESCRIPTION: prokaryotic expression vector
US-08-487-283A-18

Query Match 36.3%; Score 1199.2; DB 3; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGGACACCATCAAGTGGTCAAACTTTTCGGCGGTATGGCATGATAGCGCCGGAGAGA 60
DB 4810 CGGACACCATCAAGTGGTCAAACTTTTCGGCGGTATGGCATGATAGCGCCGGAGAGA 4751

QY 61 GTCAATTTCAGGGTGGTGAATGTGAACACAGTAACTATACGATGTGCGAGAGTATGCCG 120
DB 4750 GTCAATTTCAGGGTGGTGAATGTGAACACAGTAACTATACGATGTGCGAGAGTATGCCG 4691

QY 121 GTGTCTCTTATCAGACCGTTTCGGCGGTGGTGAACACAGGCCACGCTTTCGGGAAAA 180
DB 4690 GTGTCTCTTATCAGACCGTTTCGGCGGTGGTGAACACAGGCCACGCTTTCGGGAAAA 4631

QY 181 CGCGGMAAAGTGGAGCGCGGATGGCGGAGCTGAATTACATTCCTCAACCGCGTGGCAC 240
DB 4630 CGCGGMAAAGTGGAGCGCGGATGGCGGAGCTGAATTACATTCCTCAACCGCGTGGCAC 4571

QY 241 AACCACTGGCGGCAAAACAGTGTGCTGATTGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
DB 4570 AACCACTGGCGGCAAAACAGTGTGCTGATTGGCGTGGCCACCTCCAGTCTGGCCCTGC 4511

QY 301 ACGCGCCGTCGCAAAATTGTGCGGGCGATTAATCTGCGCGCGATCAACTGGGTGCGAGCG 360

RESULT 10

PCT-US96-05611A-21/c
; Sequence 21, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Leonardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott

Db 4510 ACGCGCGTCGCAAAATTTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCGCG 4451
QY 361 TGGTGGTGTGATGATGTAAGCGCGCTGGAAGCGCTGTAAGCGCGCGGTGCAACAATC 420
Db 4450 TGGTGGTGTGATGTAAGCGCGCTGGAAGCGCTGTAAGCGCGCGGTGCAACAATC 4391
QY 421 TTTCTCGCGCAACCGCGTCAAGTGGGCTGATCAATATATATATATATATATATATAT 480
Db 4390 TTTCTCGCGCAACCGCGTCAAGTGGGCTGATCAATATATATATATATATATATATAT 4331
QY 481 TTTCTGTGGAAGCTGCGCTGCACTAATATGTTCCGGCGTATTTCTTGTATCTCTGAC 540
Db 4330 TTTCTGTGGAAGCTGCGCTGCACTAATATGTTCCGGCGTATTTCTTGTATCTCTGAC 4271
QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGGTGAGCATC 600
Db 4270 CACCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGGTGAGCATC 4211
QY 601 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAGTCTGTCTCG 660
Db 4210 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAGTCTGTCTCG 4151
QY 661 CGCGTCTCGCTGCGCTGCGCTGGGCAATAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
Db 4150 CGCGTCTCGCTGCGCTGCGCTGGGCAATAATATCTCACTCGCAATCAAAATTCAGCGCATAG 4091
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 4090 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTGGTGGCAACGATCAGATGGCGTGGCGGCA 840
Db 4030 ATGAGGGCATCGTTCCCACTCGCATGCTGGTGGTGGCAACGATCAGATGGCGTGGCGGCA 3971
QY 841 TGGCGCCATTACCGAGTCCGGGCTGCGGCTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db 3970 TGGCGCCATTACCGAGTCCGGGCTGCGGCTGGTGGCGGATATCTCGGTAGTGGGATAG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATATCCCGCGCTTAAACCAACCATCAAAAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATATATCCCGCGCTTAAACCAACCATCAAAAGGATTTTC 3851
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db 3850 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 3791
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3790 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 3731
QY 1081 CGCAACCGCGCTCTCCCGCGCTTGGCGGATTCATTAAATGAGTGGCAGCAGGATTT 1140
Db 3730 CGCAACCGCGCTCTCCCGCGCTTGGCGGATTCATTAAATGAGTGGCAGCAGGATTT 3671
QY 1141 CCGGACTGGAAGCGGCGAGTGGAGCGCAACGCAATTAATGTGAGTGTAGTCTACTCATTAG 1200
Db 3670 CCGGACTGGAAGCGGCGAGTGGAGCGCAACGCAATTAATGTGAGTGTAGTCTACTCATTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607

Query Match	36.3%;	Score 1199.2;	DB 5;	Length 5248;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1201;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	CGGACCATCGAATGGTGC	AAACCTTTCGCGGTATG	CGATGATAGCGCCCGGAAGAGA 60
DB	4810	CGGACCATCGAATGGCG	AAACCTTTCGCGGTATG	CGATGATAGCGCCCGGAAGAGA 4751
QY	61	GTCAAATCAGGGTGGTGA	ATGTGAACACAGTAACT	TATACGATGCGCAGATATGCCG 120
DB	4750	GTCAAATCAGGGTGGTGA	ATGTGAACACAGTAACT	TATACGATGCGCAGATATGCCG 4691
QY	121	GTGTCTCTTATCAGACCG	TTTTCCGCGTGGTGAAC	CAGCGACGACCGTTTCTCGGAAA 180
DB	4690	GTGTCTCTTATCAGACCG	TTTTCCGCGTGGTGAAC	CAGCGACGACCGTTTCTCGGAAA 4631
QY	181	CGCGGGAAAAAGTGGAA	CGCGCGATGCGCGAGCT	GAAATTCACATCCCAACCGCGTGGCAC 240
DB	4630	CGCGGGAAAAAGTGGAA	CGCGCGATGCGCGAGCT	GAAATTCACATCCCAACCGCGTGGCAC 4571
QY	241	AACAACCTGGCGGGAAC	AGTGTCTGATTTGGCGT	TGGCCACCTCCAGTCTGGCCCTTGC 300
DB	4570	AACAACCTGGCGGGAAC	AGTGTCTGATTTGGCGT	TGGCCACCTCCAGTCTGGCCCTTGC 4511
QY	301	ACGCGCGGTGC	CAAAATGTTCGGCGCGA	TTAATCTTCGCGCGGATCAATGGGTGCCACGC 360

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pET-
US-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 60
DB 4622 CGGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 4563

QY 61 GTCAAATTCAGGGTGTGAATGTGAACACAGTAACCTTATACGATGTCGACAGTATGCCG 120
DB 4562 GTCAAATTCAGGGTGTGAATGTGAACACAGTAACTTATACGATGTCGACAGTATGCCG 4503

QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCACCGTTCCTCGCAAAA 180
DB 4502 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCACCGTTCCTCGCAAAA 4443

QY 181 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
DB 4442 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 4383

QY 241 AACAACTCGCGGCAACAGTGTGCTGATTGGCGTTGGCCACCTCCAGTTCGGCCCTGC 300
DB 4382 AACAACTCGCGGCAACAGTGTGCTGATTGGCGTTGGCCACCTCCAGTTCGGCCCTGC 4323

QY 301 ACGCGCGTTCGCAAAATTTGCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 4322 ACGCGCGTTCGCAAAATTTGCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 4263

QY 361 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCGCTGTAAAGCGGCGTGCACATC 420
DB 4262 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCGCTGTAAAGCGGCGTGCACATC 4203

QY 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACACGAGGATGCA 480
DB 4202 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACACGAGGATGCA 4143

QY 481 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTGACCCAGA 540
DB 4142 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTGACCCAGA 4083

QY 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCACTGGCGGTGGAGCATC 600
DB 4082 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCACTGGCGGTGGAGCATC 4023

QY 601 TGGTGCATTTGGTTCACGACCAATCGCGCTTTAGCGGGCCCATTAAGTTCTGTCGG 660
DB 4022 TGGTGCATTTGGTTCACGACCAATCGCGCTTTAGCGGGCCCATTAAGTTCTGTCGG 3963

QY 661 CCGCTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
DB 3962 CCGCTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 3903

QY 721 CGGAAACGGGAAGCGGACGTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
DB 3902 CGGAAACGGGAAGCGGACGTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 3843

QY 781 ATGAGGGCATCTTCCCACTGGATGCTGGTTGCCACGATCAGATGGCGTGGCGGCA 840
DB 3842 ATGAGGGCATCTTCCCACTGGATGCTGGTTGCCACGATCAGATGGCGTGGCGGCA 3783

QY 841 TGCAGCGCATTTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGATACG 900
DB 3782 TGCAGCGCATTTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGATACG 3723

QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGAACACCATCAAAACAGGATTTTC 960
DB 3722 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGAACACCATCAAAACAGGATTTTC 3663

QY 961 GCCTGCTGGGCAAAACACGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 3662 GCCTGCTGGGCAAAACACGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 3603

QY 1021 AGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 1080
DB 3602 AGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 3543

QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTAATTAATGCAGCTGCGACAGGTTT 1140
DB 3542 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTAATTAATGCAGCTGCGACAGGTTT 3483

QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 3482 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 3423

QY 1201 GCAC 1204
DB 3422 GCAC 3419

RESULT 12
US-08-929-967-1/c
; Sequence 1, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Ruppert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dregler, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGATGATAGCGCCCGGAGAGA 60
Db 4810 CGGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGATGATAGCGCCCGGAGAGA 4751
QY 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATGCCG 120
Db 4750 GTCATTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACACAGCGCCAGCCAGCTTTCTCGGAAAA 180
Db 4690 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACACAGCGCCAGCCAGCTTTCTCGGAAAA 4631
QY 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db 4630 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 4571
QY 241 AACACTGGCGGCAACAGTGGTGTGCTGATTTGGGTTGGCCAGTCTCCAGTCTGGCCCTGC 300
Db 4570 AACACTGGCGGCAACAGTGGTGTGCTGATTTGGGTTGGCCAGTCTCCAGTCTGGCCCTGC 4511
QY 301 AGCGCCGTTCGCAAAATTTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGGCCAGC 360
Db 4510 AGCGCCGTTCGCAAAATTTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGGCCAGC 4451
QY 361 TGGTGGTTCGATGTGAAGCAAGCGCGCTGAAGCGCTGTAAAGCGCGGTGCACAATC 420
Db 4450 TGGTGGTTCGATGTGAAGCAAGCGCGCTGAAGCGCTGTAAAGCGCGGTGCACAATC 4391
QY 421 TTCTCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGGCCAGC 480
Db 4390 TTCTCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGGCCAGC 4331
QY 481 TTGTGTGGAAGCTGCCTGCCTCAATATGTTCCGGCGTTATTTCTTGATGCTCTCAGCAGA 540
Db 4330 TTGTGTGGAAGCTGCCTGCCTCAATATGTTCCGGCGTTATTTCTTGATGCTCTCAGCAGA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGATCGGCGTGGAGCATC 600
Db 4270 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGATCGGCGTGGAGCATC 4211
QY 601 TGGTGCATTCGGTCCAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGCTCTCGG 660
Db 4210 TGGTGCATTCGGTCCAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGCTCTCGG 4151
QY 661 CGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4150 CGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
QY 721 CGGAACGGGAAGGCGACTGGAGTCCCATGTCCGGTTTTCACCAACCATGCAATGCTGA 780
Db 4090 CGGAACGGGAAGGCGACTGGAGTCCCATGTCCGGTTTTCACCAACCATGCAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTGCTGCTGTGTCACACGATCAGATGGCGCTGGGCGCAA 840
Db 4030 ATGAGGGCATCGTTCCCACTGCTGCTGTGTCACACGATCAGATGGCGCTGGGCGCAA 3971
QY 841 TGGCGGCATTCAGAGTCCGGGTGCGGTGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 3970 TGGCGGCATTCAGAGTCCGGGTGCGGTGTGGTGGGATATCTCGGTAGTGGGATACG 3911
QY 901 ACAGATCCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCATCAACAGATTTTC 960
Db 3910 ACAGATCCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCATCAACAGATTTTC 3851
QY 961 GCTGTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 3850 GCTGTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 3791
QY 1021 AGGCAATCAGTGTGGCCGCTCTCACTGGTGAAGAAAAACACCGCTGGCCGCCAATA 1080
Db 3790 AGGCAATCAGTGTGGCCGCTCTCACTGGTGAAGAAAAACACCGCTGGCCGCCAATA 3731
QY 1081 CGCAAAACCGCTCTCCCGCGCGTGGCCGATCAATTAATGCAGCTGGCAGCAGAGTTT 1140

Db 3730 CGCAAAACCGCTCTCCCGCGCGTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT 3671
QY 1141 CCCCACTGGAAAGCGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 1200
Db 3670 CCCCACTGGAAAGCGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607
RESULT 13
US-09-702-705-785/c
; Sequence 785, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Lihun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-785
Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGATGATAGCGCCCGGAGAGA 60
Db 4677 CGGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGATGATAGCGCCCGGAGAGA 4618
QY 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATGCCG 120
Db 4617 GTCATTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACACAGCGCCAGCCAGCTTTCTCGGAAAA 180
Db 4557 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACACAGCGCCAGCCAGCTTTCTCGGAAAA 4498
QY 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 4438
QY 241 AACACTGGCGGCAACAGTGGTGTGCTGATTTGGGTTGGCCAGTCTCCAGTCTGGCCCTGC 300
Db 4437 AACACTGGCGGCAACAGTGGTGTGCTGATTTGGGTTGGCCAGTCTCCAGTCTGGCCCTGC 4378
QY 301 AGCGCCGTTCGCAAAATTTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGGCAGC 360
Db 4377 AGCGCCGTTCGCAAAATTTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGGCAGC 4318
QY 361 TGGTGGTTCGATGTGAAGCAAGCGCGGTGCAGACCTGTAAAGCGCGGTGCACAATC 420
Db 4317 TGGTGGTTCGATGTGAAGCAAGCGCGGTGCAGACCTGTAAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCAACCGCTCAGTGGCTGATCAATTAATGCAGCTGGCAGCAGAGTTT 480

Db 4257 TTCTCGCGCAACGCGTCAGTGGCTGATCATTAATACTATCCGCTGGATGACCAAGATGCCA 4198
QY 481 TTGCTGTGGAAGCTGCCTGCATAATAGTTCGGCGGTATTTCTTGATGTCCTGACCAAGA 540
Db 4197 TTGCTGTGGAAGCTGCCTGCATAATAGTTCGGCGGTATTTCTTGATGTCCTGACCAAGA 4138
QY 541 CACCCATCAACAGTATATTTCTCCCATGAAGCGGTACGACCTGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATATTTCTCCCATGAAGCGGTACGACCTGGCGGTGGAGCATC 4078
QY 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGATTCTGTCTCGG 660
Db 4077 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGATTCTGTCTCGG 4018
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGGCGACCTGGAGTGCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACCTGGAGTGCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 3898
QY 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAACGATGCGCTGGGCGCAA 840
Db 3897 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAACGATGCGCTGGGCGCAA 3838
QY 841 TCGCGCCATTAACGAGTCCGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCATTAACGAGTCCGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 3778
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAACGAGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAACGAGATTTTC 3718
QY 961 GCCTCTGCGGCAACACGCGTGGACCGCTTCTGCTCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 3717 GCCTCTGCGGCAACACGCGTGGACCGCTTCTGCTCAACTCTCTCAGGCGCAGGCGGTGA 3658
QY 1021 AGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3657 AGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 3598
QY 1081 CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTATGCACTGGCAGCAGCAGAGTTT 1140
Db 3597 CGCAACCGCTCTCTCCCGCGGCTTGGCGGATTCATTATGCACTGGCAGCAGCAGAGTTT 3538
QY 1141 CCGCACTGGAAGCGGCGAGTCAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 3537 CCGCACTGGAAGCGGCGAGTCAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 14

US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Panger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mammion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGGAAGAGA 60
Db 4677 CCGACACCATCGAATGGTGGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGGAAGAGA 4618
QY 61 GTCAATTCAGGGTGGTGAATGTAACAGTAAGCTTATACGATGTCGACGATGTCGG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTAACAGTAAGCTTATACGATGTCGACGATGTCGG 4558
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACAGGCCAGCCACGTTTCTCGAAAA 180
Db 4557 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACAGGCCAGCCACGTTTCTCGAAAA 4498
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAACAGTCTGTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAACAGTCTGTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 4378
QY 301 ACGGCGCGTCCGCAAAATTTCTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGGCAGCG 360
Db 4377 ACGGCGCGTCCGCAAAATTTCTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGGCAGCG 4318
QY 361 TGGTGGTTCGATGGTAGAACGAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 4317 TGGTGGTTCGATGGTAGAACGAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 4258
QY 421 TTCTCGCGCAACGCGTCACTGGGTGATCAATTAATCTCGCTGGATGACCAAGGATGCA 480
Db 4257 TTCTCGCGCAACGCGTCACTGGGTGATCAATTAATCTCGCTGGATGACCAAGGATGCA 4198
QY 481 TTGTCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCAAGA 540
Db 4197 TTGTCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCAAGA 4138
QY 541 CACCCATCAACAGTATATTTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATATTTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATC 4078
QY 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGATTCTGTCTCGG 660
Db 4077 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGATTCTGTCTCGG 4018
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGGCGACCTGGAGTGCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACCTGGAGTGCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 3898
QY 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAACGATGCGCTGGGCGCAA 840
Db 3897 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAACGATGCGCTGGGCGCAA 3838
QY 841 TCGCGCCATTAACGAGTCCGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCATTAACGAGTCCGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 3778


```
QY 901 ACATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACACCATCAAAACAGGATTTTC 960
Db 3777 ACATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACACCATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAACACAGCGTGGACCGTTGCTGCAACTCTCTCAGGGCCACGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAACACAGCGTGGACCGTTGCTGCAACTCTCTCAGGGCCACGCGGTGA 3658
QY 1021 AGGCAATCAGCTGTGTCGCCGCTCTCACTGGTGAAAGAAACACACCTGGCGCCCAATA 1080
Db 3657 AGGCAATCAGCTGTGTCGCCGCTCTCACTGGTGAAAGAAACACACCTGGCGCCCAATA 3598
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGAATTCATTAATGAGCTGGCAGACAGGTTT 1140
Db 3597 CGCAACCGCTCTCCCGCGGTTGGCGAATTCATTAATGAGCTGGCAGACAGGTTT 3538
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTACCTCACTCATTTAG 1200
Db 3537 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTACCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474
```

RESULT 15

```
US-09-614-124B-785/c
; Sequence 785, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614.124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1688
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-785
```

```
Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGATGTGTGCAAAACCTTTCGCGTATGGATGATAGCGCCCGAAGAGA 60
Db 4677 CGGACCAATCGATGTGTGCAAAACCTTTCGCGTATGGATGATAGCGCCCGAAGAGA 4618
QY 61 GTCAATTCAGGGTGGTGAATGTGAACAGTAACTTATACGATGTGCGAGAGTATGCCG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTGAACAGTAACTTATACGATGTGCGAGAGTATGCCG 4558
QY 121 GTGTCTTTATCAGACCGTTTCCCGTGTGTAACAGGCGAGCCAGCTTTCTGCGAAA 180
Db 4557 GTGTCTTTATCAGACCGTTTCCCGTGTGTAACAGGCGAGCTTTCTGCGAAA 4498
QY 181 CGCGGAAAAGTGAAGCGGCGATGCGAGCTGAATTAATTCACCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAGTGAAGCGGCGATGCGAGCTGAATTAATTCACCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCGCACCTCCAGCTCTGGCCCTGC 300
```

```
Db 4437 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCGCACCTCCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGTTCGCAAAATTTGTTCGCGCGATTAATCTCGCGCGGATCAAACTGGGTGCCAGCG 360
Db 4377 ACGCGCGTTCGCAAAATTTGTTCGCGCGATTAATCTCGCGCGGATCAAACTGGGTGCCAGCG 4318
QY 361 TGGTGGTGTGATGGTGAAGCAAGCGCGTTCGAGCTGTAAAGCGCGGTGCACAATC 420
Db 4317 TGGTGGTGTGATGGTGAAGCAAGCGCGTTCGAGCTGTAAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGACAGGATGCCA 480
Db 4257 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGACAGGATGCCA 4198
QY 481 TTGCTGTGGAAGCTGCTGCACATAATTTCCGCGTTATTTCTTGATGTCTCTGACGAGA 540
Db 4197 TTGCTGTGGAAGCTGCTGCACATAATTTCCGCGTTATTTCTTGATGTCTCTGACGAGA 4138
QY 541 CACCATCAACAGCTATTTATTTCTCCCATGAAGCGGTAGCGACTGGGCGTGGAGATC 600
Db 4137 CACCATCAACAGCTATTTATTTCTCCCATGAAGCGGTAGCGACTGGGCGTGGAGATC 4078
QY 601 TGGTTCGCAATTTGGGTCAACAGCAAAATCGCGCTTTAGCGGGCCCAATTAAGTTCTGTCGG 660
Db 4077 TGGTTCGCAATTTGGGTCAACAGCAAAATCGCGCTTTAGCGGGCCCAATTAAGTTCTGTCGG 4018
QY 661 CGCGTCTGCGTCTGGCTGGCTGGCGATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGGCTGGCTGGCGATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTCAACAAACCATGCAATCTGA 780
Db 3957 CGGAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTCAACAAACCATGCAATCTGA 3898
QY 781 ATGAGGCGATTCGTTCCACATCGCATGCTGTTGCCAAGCATCAGATGGGCGTGGGCGCAA 840
Db 3897 ATGAGGCGATTCGTTCCACATCGCATGCTGTTGCCAAGCATCAGATGGGCGTGGGCGCAA 3838
QY 841 TGGCGCCCATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 3837 TGGCGCCCATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 3778
QY 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTTAAACACCATCAAAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTTAAACACCATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 3658
QY 1021 AGGGCAATCAGCTGTGTCGCCGCTCTCACTGGTGAAAGAAACACACCTGGCGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTGTCGCCGCTCTCACTGGTGAAAGAAACACACCTGGCGCCCAATA 3598
QY 1081 CGCAAAACCGCTCTCCCGCGGTTTGGCGAATTCATTAATGAGCTGGCAGCAGCAGGTTT 1140
Db 3597 CGCAAAACCGCTCTCCCGCGGTTTGGCGAATTCATTAATGAGCTGGCAGCAGCAGGTTT 3538
QY 1141 CCGGACTGAAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTACCTCACTCATTTAG 1200
Db 3537 CCGGACTGAAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTACCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474
```

Search completed: August 3, 2005, 02:32:27
Job time : 357.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 14:24:04 ; Search time 1094.6 Seconds
(without alignments)
17846.836 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacacacgaatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	3300	100.0	3300	4	AAD11590	Partial s
2	3216.8	97.5	3300	4	AAD11592	Partial s
3	3207.2	97.2	3300	4	AAD11588	Partial s
4	3207.2	97.2	3300	4	AAD11589	Partial s
5	3207.2	97.2	3300	4	AAD11591	Partial s
6	2705.8	82.0	8101	4	AAD25681	Nucleotid
7	2705.8	82.0	8101	6	AB153238	Nucleotid
8	2690.4	81.5	7259	12	ADO23608	DNA encod
9	2690.4	81.5	7322	12	ADO23613	DNA encod
10	2690.4	81.5	7352	12	ADO23603	DNA encod
11	2690.4	81.5	7370	12	ADO23598	DNA encod
12	2690.4	81.5	7370	12	ADO23639	DNA encod
13	2690.4	81.5	7370	12	ADO23649	DNA encod
14	2690.4	81.5	7370	12	ADO23644	DNA encod
15	2690.4	81.5	7403	12	ADO23594	DNA encod
16	2690.4	81.5	7442	12	ADO23590	DNA encod
17	2690.4	81.5	7478	12	ADO23584	DNA encod
18	2688.8	81.5	7112	12	ADO23618	DNA encod
19	2636.8	79.9	6806	6	AB149925	Maltose b
20	2605.4	79.0	7553	12	ADO23588	DNA encod

21	1609.4	48.8	5558	12	ADL90419	Adl90419 Clostridi
22	1474	44.7	4700	9	ACF06053	Act06053 Vector pJ
23	1472.4	44.6	4700	12	ADL18670	Adl18670 Vector pJ
24	1462.4	44.3	4920	10	ADH73599	Adh73599 Plasmid p
25	1462.4	44.3	4935	10	ADH73598	Adh73598 Plasmid p
26	1462.4	44.3	4945	10	ADH73600	Adh73600 Plasmid p
27	1462.4	44.3	4951	10	ADH73601	Adh73601 Plasmid p
28	1418.2	43.0	5903	8	ABZ23939	Abz23939 Nucleotid
29	1417.6	43.0	1511	6	ABQ73200	Abq73200 E. coli t
30	1417.6	43.0	5926	2	AAV32977	Aav32977 Tn7 donor
31	1417.6	43.0	5926	6	AAD45059	Aad45059 Transposo
32	1417.6	43.0	5926	12	ADG46817	Adg46817 Donor pla
33	1362	41.3	5201	12	ADL72229	Adl72229 DNA seque
34	1362	41.3	5201	12	ADL72228	Adl72228 DNA seque
35	1238.4	37.5	5024	9	ACF06051	Act06051 Plasmid p
36	1238.4	37.5	5024	12	ADL18668	Adl18668 Plasmid p
37	1237.2	37.5	1922	6	ABQ73089	Abq73089 MBP and z
38	1237.2	37.5	1922	10	ADF83622	Adf83622 MBP-human
39	1204	36.5	4557	2	AAT90491	Aat90491 Vector pl
40	1204	36.5	4969	12	ADO33522	Ado33522 Expressio
41	1204	36.5	5006	10	ADE24112	Ade24112 plasmid p
42	1204	36.5	5064	12	ADO33532	Ado33532 Vector pG
43	1204	36.5	5082	12	ADO33531	Ado33531 Vector pG
44	1204	36.5	6259	6	AAD29720	Aad29720 Plasmid p
45	1204	36.5	6823	3	AAC55459	Aac55459 Destinati

ALIGNMENTS

RESULT 1		AAD11590	
ID	AAD11590 standard; DNA; 3300 BP.		
XX	AAD11590;		
XX			
DT	24-SEP-2001 (first entry)		
XX			
DE	Partial sequence of pMal-m3 and 2fPm3 DNA.		
XX			
KW	Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;		
KW	modulation; plant technology; agriculture; ds.		
XX	Unidentified.		
OS			
XX			
PH	Key	Location/Qualifiers	
FT	CDS	2719..3270	
FT		/*tag= a	
FT		/product= "ZFPm3 protein"	
FT		/note= "CDS does not include start and stop codon"	
FT		/partial	
FT	primer_bind	2740..2790	
FT		/*tag= b	
FT		/bound_moiety= "F1-f2 primer"	
FT	primer_bind	2770..2850	
FT		/*tag= c	
FT		/bound_moiety= "F1-f1 primer"	
FT	primer_bind	2824..2889	
FT		/*tag= d	
FT		/bound_moiety= "F2-b primer"	
FT	primer_bind	2867..2940	
FT		/*tag= e	
FT		/bound_moiety= "F2-f primer"	
FT	primer_bind	2916..2973	
FT		/*tag= f	
FT		/bound_moiety= "F3-b1 primer"	
FT	primer_bind	2953..3021	
FT		/*tag= g	
FT		/bound_moiety= "F3-b2 primer"	
FT	primer_bind	2992..3042	
FT		/*tag= h	
FT		/bound_moiety= "F4-f2 primer"	
FT	primer_bind	3022..3102	

```
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"
XX
PN WO200152620-A2.
PD
PD 26-JUL-2001..
XX
PF 19-JAN-2001; 2001WO-US001817.
XX
XX
PR 21-JAN-2000; 2000US-0177468P.
PR 21-JUL-2000; 2000US-00620897.
XX
PA (SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI Barbas CF, Stege JT, Guan X, Dalmia B;
XX WPI; 2001-465325/50.
DR P-PSDB; AAE06003.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 143-145; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence. Where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m3 and ZFPm3 DNA
XX
SQ Sequence 3300 BP; 845 A; 878 C; 877 G; 700 T; 0 U; 0 Other;
Query Match 100.0%; Score 3300; DB 4; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTATACGATGTGCGAGAGTATGCCG 120
DB 61 GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGTTCCTCGGAAA 180
DB 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGTTCCTCGGAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCACCAACCGCGTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCACCAACCGCGTGGCAC 240
```

```
DB 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCACCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCCGTTCGCAAAATTTGTTCGCGCGATTAATCTTCGCGCGGATCAACTGGTTCGCCAGC 360
DB 301 ACGCCCGTTCGCAAAATTTGTTCGCGCGATTAATCTTCGCGCGGATCAACTGGTTCGCCAGC 360
QY 361 TGTGTGTTCGATGTAGAACGAAAGCGCGCTGCGAAGCTGTGTAAGCGCGGTCACAATC 420
DB 361 TGTGTGTTCGATGTAGAACGAAAGCGCGCTGCGAAGCTGTGTAAGCGCGGTCACAATC 420
QY 421 TTCTCGCGCAACCGGTCAAGTGGGCTGATCATTAATCTTCGCGCGGATTAATCTTCGCGCGG 480
DB 421 TTCTCGCGCAACCGGTCAAGTGGGCTGATCATTAATCTTCGCGCGGATTAATCTTCGCGCGG 480
QY 481 TTGCTGTGGAAGCTGCCTGCACCTAATGTTCCGGGGTTAATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGGAAGCTGCCTGCACCTAATGTTCCGGGGTTAATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGACTGGCGTGGAGCATC 600
DB 541 CACCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGACTGGCGTGGAGCATC 600
QY 601 TGTTCGCAATGGGTACCCAGCAATCGCGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG 660
DB 601 TGTTCGCAATGGGTACCCAGCAATCGCGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCGTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCGTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGCGGACGTGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAAACGGGAAGCGGACGTGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCGATCTGTTCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCAA 840
DB 781 ATGAGGCGATCTGTTCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCAA 840
QY 841 TGGCGGCCATTACCGAGTCCGGCTCGCGTTCGGTTCGGATATCTCGTAGTGGGATACG 900
DB 841 TGGCGGCCATTACCGAGTCCGGCTCGCGTTCGGTTCGGATATCTCGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAAACCCCTTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAAACCCCTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGGACGACGAGTTT 1140
DB 1081 CGCAAAACCCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGGACGACGAGTTT 1140
QY 1141 CCGCATGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG 1200
DB 1141 CCGCATGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG 1200
QY 1201 GCAGAACTCTATGTTTGCAGCTTATCATGCACTGCGAGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCAGAACTCTATGTTTGCAGCTTATCATGCACTGCGAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGACCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATAATCTGCAATAATTCG 1320
DB 1261 TCAGGACGACCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATAATCTGCAATAATTCG 1320
```

QY	1321	TGTCGCTCAAGCGCAGCTCCCGTTCCTGGATAATGTTTTTGGCCGACATCATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCAGCTCCCGTTCCTGGATAATGTTTTTGGCCGACATCATACGGTT	1380
QY	1381	CTGGCAAAATATTTCTGAAATGAGCTGTCGACAAATTAATCATCGGCTCGTATAATGTTGTGA	1440
Db	1381	CTGGCAAAATATTTCTGAAATGAGCTGTCGACAAATTAATCATCGGCTCGTATAATGTTGTGA	1440
QY	1441	ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTTAGTGTTTTTCACGA	1500
Db	1441	ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTTAGTGTTTTTCACGA	1500
QY	1501	GCATTTCAACACAGGACCATAGATTATGAAAACTGAAAGGTAATAACTGTTATCTGG	1560
Db	1501	GCATTTCAACACAGGACCATAGATTATGAAAACTGAAAGGTAATAACTGTTATCTGG	1560
QY	1561	ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGAGAGAAATTCGCCAGGTT	1680
Db	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGAGAGAAATTCGCCAGGTT	1680
QY	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGCGTAC	1740
Db	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGCGTAC	1740
QY	1741	GCTCAATCTGGGCTGTGGCTGAAATCAACCCGGACAAGCGTTCCAGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGGGCTGTGGCTGAAATCAACCCGGACAAGCGTTCCAGACAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGATGCGCTAGCTTACAAAGCGAAGCTGATTGCTTACCCGATCGCTGT	1860
Db	1801	CCGTTTACCTGGATGCGCTAGCTTACAAAGCGAAGCTGATTGCTTACCCGATCGCTGT	1860
QY	1861	GAAGGTTATCGCTGATTTAACAAGATCTGCTGCGGAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAAGGTTATCGCTGATTTAACAAGATCTGCTGCGGAACCCGCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGACGGGGTTATCGCTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGACGGGGTTATCGCTTCAAG	2040
QY	2041	TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG	2100
QY	2101	GCTCTGACCTTCTGGTTGACCTGATTAATAACAACACATGAATGCAGACCCGATTAC	2160
Db	2101	GCTCTGACCTTCTGGTTGACCTGATTAATAACAACACATGAATGCAGACCCGATTAC	2160
QY	2161	TCCATCGCAAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
QY	2221	GCATGGTCAACATCGACACCGAAGCTGTTGGCTGCTGAGCGGAGGTATTAAACCGCCAGT	2280
Db	2221	GCATGGTCAACATCGACACCGAAGCTGTTGGCTGCTGAGCGGAGGTATTAAACCGCCAGT	2280
QY	2281	ANGGTCACCATCCAAACCGTTGTTGGCTGCTGAGCGGAGGTATTAAACCGCCAGT	2340
Db	2281	ANGGTCACCATCCAAACCGTTGTTGGCTGCTGAGCGGAGGTATTAAACCGCCAGT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTACTGATGAAGTCTG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTACTGATGAAGTCTG	2400

RESULT 2

AAD11592

ID AAD11592 standard; DNA; 3300 BP.

XX AAD11592;

AC AAD11592;

XX 24-SEP-2001 (first entry)

DT Partial sequence of pMal-Ap3 and ZFPap3 DNA.

XX Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;

KW modulation; plant technology; agriculture; Ap3; APETALA3; ds.

QY	2401	GAAGCGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAACTCTTACGAGGAAG	2460
Db	2401	GAAGCGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAACTCTTACGAGGAAG	2460
QY	2461	TTGGCGAAAGATCCACGTTATTTGGCGCCACCATGAAAAACGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTTATTTGGCGCCACCATGAAAAACGCCAGAAAGGTGAATCATG	2520
QY	2521	CCGAACATCCCGAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
QY	2581	GCCAGCGTCTGAGCTGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGTCTGAGCTGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG	2640
QY	2641	AACAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC	2700
Db	2641	AACAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC	2700
QY	2701	GGATCCTCTTCTGTGGCCGACGCGGCTTCGAGCCCGGGGAGAACCTTATGCTTGT	2760
Db	2701	GGATCCTCTTCTGTGGCCGACGCGGCTTCGAGCCCGGGGAGAACCTTATGCTTGT	2760
QY	2761	CGGNAATGCTGTAAGTCTTTCAGCGATCCTGCGCACCTGTTCCGCCACAGCGTACCCAC	2820
Db	2761	CGGNAATGCTGTAAGTCTTTCAGCGATCCTGCGCACCTGTTCCGCCACAGCGTACCCAC	2820
QY	2821	ACGGTGAAAAACCGTATAAATGCCAGATGTCGGCAAACTCTTTAGCACACAGCGGCTCC	2880
Db	2821	ACGGTGAAAAACCGTATAAATGCCAGATGTCGGCAAACTCTTTAGCACACAGCGGCTCC	2880
QY	2881	CTGGTGCCCATCAACCGCACTCATCTGGCGAAGGCCATACAAATGTCCAGATGTGGC	2940
Db	2881	CTGGTGCCCATCAACCGCACTCATCTGGCGAAGGCCATACAAATGTCCAGATGTGGC	2940
QY	2941	AAGTCTTTTCAGCCAGAGCTCCAGCTGTCGCGCACCAACCGTACTCACACGGGAGAG	3000
Db	2941	AAGTCTTTTCAGCCAGAGCTCCAGCTGTCGCGCACCAACCGTACTCACACGGGAGAG	3000
QY	3001	CCCTATGCTTGTCCGGAATGTGTAAGTCTTTCAGCCAGAGCAGCTCCCTGGTGGCCAC	3060
Db	3001	CCCTATGCTTGTCCGGAATGTGTAAGTCTTTCAGCCAGAGCAGCTCCCTGGTGGCCAC	3060
QY	3061	CAGCGTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGT	3120
Db	3061	CAGCGTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGT	3120
QY	3121	GACTGCCCGGACCTTGCTCGCCATCAACGCACTCATCTGCGGAGAACCATACAAATGT	3180
Db	3121	GACTGCCCGGACCTTGCTCGCCATCAACGCACTCATCTGCGGAGAACCATACAAATGT	3180
QY	3181	CCAGATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTCGTCCGGCACCAACGTTACTAC	3240
Db	3181	CCAGATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTCGTCCGGCACCAACGTTACTAC	3240
QY	3241	ACCGGTAAAAAATAGTGGCCAGCGCGCCAGTACCGGTACGAGCTTCCGACTACGCT	3300
Db	3241	ACCGGTAAAAAATAGTGGCCAGCGCGCCAGTACCGGTACGAGCTTCCGACTACGCT	3300

```

XX OS Unidentified.
XX OS
XX PH Location/Qualifiers
XX FT 2719..3270
XX FT /tag= a
XX FT /product= "ZFPp3 protein"
XX FT /transl_except= (pos:3124..3126, aa:Ser)
XX FT /transl_except= (pos:3127..3129, aa:Ser)
XX FT /transl_except= (pos:3136..3138, aa:Val)
XX FT /note= "CDS does not include start and stop codon"
XX FT /partial
XX FT 2740..2790
XX FT /tag= b
XX FT /bound_moiety= "F1-f2 primer"
XX FT 2770..2850
XX FT /tag= c
XX FT /bound_moiety= "F1-f1 primer"
XX FT 2824..2889
XX FT /tag= d
XX FT /bound_moiety= "F2-b primer"
XX FT 2867..2940
XX FT /tag= e
XX FT /bound_moiety= "F2-f primer"
XX FT 2916..2973
XX FT /tag= f
XX FT /bound_moiety= "F3-b1 primer"
XX FT 2953..3021
XX FT /tag= g
XX FT /bound_moiety= "F3-b2 primer"
XX FT 2992..3042
XX FT /tag= h
XX FT /bound_moiety= "F4-f2 primer"
XX FT 3022..3102
XX FT /tag= i
XX FT /bound_moiety= "F4-f1 primer"
XX FT 3076..3141
XX FT /tag= j
XX FT /bound_moiety= "F5-b primer"
XX FT 3119..3192
XX FT /tag= k
XX FT /bound_moiety= "F5-f primer"
XX FT 3168..3225
XX FT /tag= l
XX FT /bound_moiety= "F6-b1 primer"
XX FT 3205..3273
XX FT /tag= m
XX FT /bound_moiety= "F6-b2 primer"
XX MO200152620-A2.
XX
XX PD 26-JUL-2001.
XX
XX PF 19-JAN-2001; 2001WO-US001817.
XX
XX PR 21-JAN-2000; 2000US-0177468P.
XX PR 21-JUL-2000; 2000US-00620897.
XX
XX PA (SCRI ) SCRIPPS RES INST.
XX PA (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX PI Barbas CF, Stage JT, Guan X, Dalmia B;
XX
XX DR WPI; 2001-45325/50.
XX DR P-PSDB; AAE06005.
XX
XX PT New zinc finger proteins, useful for modulating or regulating gene
XX PT expression and metabolic pathways in plants, e.g. for treating in the
XX PT plant cells a disorder that is associated with abnormal expression of the
XX PT target gene.
XX PS Example 4; Page 148-149; 156pp; English.
XX

```

```

CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-Ap3 (APETALA3) and ZFPp3 DNA
XX
XX SQ Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;
XX
Query Match 97.5%; Score 3216.8; DB 4; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTTCAGGGTGGTGAATGTCAAAACAGTATACGTTATACGATGTGCGAGAGTATCCG 120
DB 61 GTCAATTTCAGGGTGGTGAATGTCAAAACAGTATACGTTATACGATGTGCGAGAGTATCCG 120
QY 121 GTGTCTCTTATACAGCGTTTCCCGGTGTGAACAGGCCAGCCAGCTTTCGCGAAAA 180
DB 121 GTGTCTCTTATACAGCGTTTCCCGGTGTGAACAGGCCAGCCAGCTTTCGCGAAAA 180
QY 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTCATTCCTCAACCGGTGGCAC 240
DB 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTCATTCCTCAACCGGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTCTGCTGATTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAACAGTCTGCTGATTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTGCAGAAATTTGCGCGCGATTAATTCCTCGCGCGATCAACTGGGTGCAGCG 360
DB 301 ACGCGCGTGCAGAAATTTGCGCGCGATTAATTCCTCGCGCGATCAACTGGGTGCAGCG 360
QY 361 TGTGTGTGTGATGTAGAACGAGCGCTGTAAGCGCTGTAAGCGCGGTGCACAATC 420
DB 361 TGTGTGTGTGATGTAGAACGAGCGCTGTAAGCGCTGTAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGGTGAGTGGCTGATCAATTAACCTATCCGCTGGATGACGAGATGCCA 480
DB 421 TTCTCGCGCAACCGGTGAGTGGCTGATCAATTAACCTATCCGCTGGATGACGAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTTTCGCGCGTTATTTCTTGATGTCTTGACACAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACTAATTTTCGCGCGTTATTTCTTGATGTCTTGACACAGA 540
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGCTGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGCTGGCGTGGAGCATC 600
QY 601 TGTGTGCATTGGGTACACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCGTCTCGG 660
DB 601 TGTGTGCATTGGGTACACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCGTCTCGG 660
QY 661 CGGCTCTGCTCTGGCTGGCTGAATAATCTCACTCGCAATCAAAATTCAGCCCATAG 720
DB 661 CGGCTCTGCTCTGGCTGGCTGAATAATCTCACTCGCAATCAAAATTCAGCCCATAG 720
QY 721 CCGAAACGGGAAGCGGAGTGGAGTGCATGTCGGTGTTCACAAACCAATGCAAAATGCTGA 780
DB 721 CCGAAACGGGAAGCGGAGTGGAGTGCATGTCGGTGTTCACAAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCCACTGCTGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840

```

||||| 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTGTGCGCAACGATCAGATGCGCTGGCGCGCAA 840
QY TCGCGCCCATTTACCAAGTCCCGGCTCGCGGTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db TCGCGCCCATTTACCGAGTCCGGCTCGCGGTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACACACATCAAAACAGGATTTTC 960
Db ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACACACATCAAAACAGGATTTTC 960
QY GCCTGCTGGGCAACACAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db GCCTGCTGGGCAACACAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY AGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db AGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY CGCAACCGCCTCTCCCGCGGTGGCGGATTCATTAATGCACTGGCAGCGACACAGGTTT 1140
Db CGCAACCGCCTCTCCCGCGGTGGCGGATTCATTAATGCACTGGCAGCGACACAGGTTT 1140
QY CCCGACTCGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db CCCGACTCGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTCGCAACCAATGAGTTAGCTCACTCATTTAG 1260
Db GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTCGCAACCAATGAGTTAGCTCACTCATTTAG 1260
QY TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGTGTGCAAGTGTGCAATCACTGCAATTTTCG 1320
Db TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGTGTGCAAGTGTGCAATCACTGCAATTTTCG 1320
QY TGTGCTCAAGCGCACTCCCGTTCTCGGTAATGTTTTTGGCGCGACATCATACGGTT 1380
Db TGTGCTCAAGCGCACTCCCGTTCTCGGTAATGTTTTTGGCGCGACATCATACGGTT 1380
QY CTGGCAAAATTTCTGAAATGAGCTGTGCACAATTAATCATCGGCTCGTAAATGTTGTGA 1440
Db CTGGCAAAATTTCTGAAATGAGCTGTGCACAATTAATCATCGGCTCGTAAATGTTGTGA 1440
QY ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTAGTGTGTTTTCAGGA 1500
Db ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTAGTGTGTTTTCAGGA 1500
QY GCACCTTCCCAACAGGACCATAGATTATGAAACTGGAAGGTAAGTGTGTAATCTGG 1560
Db GCACCTTCCCAACAGGACCATAGATTATGAAACTGGAAGGTAAGTGTGTAATCTGG 1560
QY ATTAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db ATTAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY ACCGAAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
Db ACCGAAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
QY GCGGCAATGCGGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740
Db GCGGCAATGCGGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740
QY GCTCAATCTGGGCTGTGGCTGAAATCAACCGGCAAAAGCGTTCCAGSACAAGCTGTAT 1800
Db GCTCAATCTGGGCTGTGGCTGAAATCAACCGGCAAAAGCGTTCCAGSACAAGCTGTAT 1800
QY CGGTTTACCTGGGATGCGGTACGTTTACACAGGCAAGCTGATGTTTACCCGATCGCTGTT 1860
Db CGGTTTACCTGGGATGCGGTACGTTTACACAGGCAAGCTGATGTTTACCCGATCGCTGTT 1860
QY GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCGGAACCCCGCAAAAACCTGGGAA 1920
Db GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCGGAACCCCGCAAAAACCTGGGAA 1920

Db 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCGGAACCCCGCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTGAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTGAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACGTTACTTACCTGCGCGCTGATGCTGCTGACGGGGTTATGCGGTTCAAG 2040
Db 1981 CTGCAAGAACGTTACTTACCTGCGCGCTGATGCTGCTGACGGGGTTATGCGGTTCAAG 2040
QY 2041 TATCAAAAACGCAAGTACGACATTTAAAGACCGTGGCGTGGATAACGCTGGCGCGAAACGG 2100
Db 2041 TATCAAAAACGCAAGTACGACATTTAAAGACCGTGGCGTGGATAACGCTGGCGCGAAACGG 2100
QY 2101 GGTCTGACCTTCTGCTGGTGAACCTGATTTAAAAACAAAACACATGAATGACAGACCGGATTAAC 2160
Db 2101 GGTCTGACCTTCTGCTGGTGAACCTGATTTAAAAACAAAACACATGAATGACAGACCGGATTAAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAATAAAGCGCAACACGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAATAAAGCGCAACACGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACCAAGTGAATTTATGGTGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACCAAGTGAATTTATGGTGTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTCTTACGAGGAAG 2460
QY 2461 TTGCGAAAGATCCACGTTATTTGCCGCCACCAATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGAAAGATCCACGTTATTTGCCGCCACCAATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGACAGTGTGATGAAGCGCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGACAGTGTGATGAAGCGCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATTAACAAACCAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAAACAAATTAACAAACCAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
QY 2701 GGATCTCTTCTGTCGGCCAGCGGCCCTCGAGCCCGGGGAGAACCCCTATCTGTT 2760
Db 2701 GGATCTCTTCTGTCGGCCAGCGGCCCTCGAGCCCGGGGAGAACCCCTATCTGTT 2760
QY 2761 CCGGAATGTTGTAAGTCTTCCAGCGATCTGCGCACCTGGTTCGCGCACAGCGGTACCCAC 2820
Db 2761 CCGGAATGTTGTAAGTCTTCCAGCGAGAGTCTTCCGCGCACAGCGGTACCCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCCCAGAGTGGGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCCAGAGTGGGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCATCAACCGACTCATCTGCGAGGAAGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACCGACTCATCTGCGAGGAAGCCATACAAATGTCCAGAAATGTGGC 2940
QY 2941 AAGTCTTTTACGCCAGCTCCAGCGCTGTCGCGCACCAACGCTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTTACGCCAGCTCCAGCGCTGTCGCGCACCAACGCTACTCACACCGGGGAGAG 3000

QY 3001 CCCTATGCTGTCCGGAATGGTAAAGTCTTTCAGCCAGAGCAGCTCCCTGGTGGCCAC 3060
Dd |||||||
3001 CCCTATGCTGTCCGGAATGGTAAAGTCTTTCAGCCAGAGCAGCTCCCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACGGGTGAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTAGT 3120
Dd |||||||
3061 CAGCGTACCCACACGGGTGAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTAGC 3120
QY 3121 GACTGCCCGACCTTGTCTCGCCATCAAGCCACTATCTACTGGCGAGAGCCATACAAATGT 3180
Dd |||||||
3121 CAGCGGCCACCTGGAAGCCATCAAGCCACTATCTACTGGCGAGAGCCATACAAATGT 3180
QY 3181 CCAGATGTGGCAAGTCTTTCTCCAACTCCAGCCATCTCGTCCGGCACCAACGTACTCAC 3240
Dd |||||||
3181 CCAGATGTGGCAAGTCTTTCTCAACTTCAGGCAACTTGGTCCGTCCACCAACGTACTCAC 3240
QY 3241 ACCGTTAAAAAATAGTGGCCAGCGCGCCAGTACCGTACGAGTTCGGACTAGCT 3300
Dd |||||||
3241 ACCGTTAAAAAATAGTGGCCAGCGCGCCAGTACCGTACGAGTTCGGACTAGCT 3300

RESULT 3

AAD11588
ID AAD11588 standard; DNA; 3300 BP.

AC AAD11588;

DT 24-SEP-2001 (first entry)

DE Partial sequence of pMal-m1 and ZFPm1 DNA.

KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.

OS Unidentified.

PH Key

FT Location/Qualifiers

FT CDS 2719..3270
FT /*tag= a
FT /product= "ZFPm1 protein"
FT /transl_except= (pos:2785..2787, aa:Asp)
FT /transl_except= (pos:2788..2790, aa:Pro)
FT /transl_except= (pos:2791..2793, aa:Gly)
FT /transl_except= (pos:2874..2874, aa:Arg)
FT /transl_except= (pos:2875..2877, aa:Ala)
FT /transl_except= (pos:2878..2880, aa:His)
FT /transl_except= (pos:2884..2886, aa:Glu)
FT /transl_except= (pos:2953..2955, aa:Gln)
FT /transl_except= (pos:2959..2961, aa:Ser)
FT /transl_except= (pos:3121..3123, aa:Arg)
FT /transl_except= (pos:3124..3126, aa:Ser)
FT /transl_except= (pos:3127..3129, aa:Asp)
FT /transl_except= (pos:3130..3132, aa:Asn)
FT /transl_except= (pos:3136..3138, aa:Val)
FT /transl_except= (pos:3139..3141, aa:Arg)
FT /transl_except= (pos:3205..3207, aa:Gln)
FT /transl_except= (pos:3208..3210, aa:Ala)
FT /transl_except= (pos:3211..3213, aa:Gly)
FT /transl_except= (pos:3214..3216, aa:His)
FT /transl_except= (pos:3220..3222, aa:Ala)
FT /transl_except= (pos:3223..3225, aa:Ser)
FT /note= "CDS does not include start and stop codon"

FT primer_bind

FT 2740..2790

FT /*tag= b

FT /bound_moiety= "F1-f2 primer"

FT 2770..2850

FT /*tag= c

FT /bound_moiety= "F1-f1 primer"

FT 2824..2889

FT /*tag= d

FT /bound_moiety= "F2-b primer"

FT primer_bind

FT 2867..2940

FT /*tag= e

FT /bound_moiety= "F2-f primer"

FT 2916..2973

FT /*tag= f

FT /bound_moiety= "F3-b1 primer"

FT 2953..3021

FT /*tag= g

FT /bound_moiety= "F3-b2 primer"

FT 2992..3042

FT /*tag= h

FT /bound_moiety= "F4-f2 primer"

FT 3022..3102

FT /*tag= i

FT /bound_moiety= "F4-f1 primer"

FT 3076..3141

FT /*tag= j

FT /bound_moiety= "F5-b primer"

FT 3119..3192

FT /*tag= k

FT /bound_moiety= "F5-f primer"

FT 3188..3225

FT /*tag= l

FT /bound_moiety= "F6-b1 primer"

FT 3205..3273

FT /*tag= m

FT /bound_moiety= "F6-b2 primer"

XX WO200152620-A2.

XX 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US001817.

XX 21-JAN-2000; 2000US-0177468P.

XX 21-JUL-2000; 2000US-00620897.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA AGRIC DISCOVERY INC.

XX Barbas CF, Stege JT, Guan X, Dalmia B;

XX WPI; 2001-465325/50.

XX P-PSDB; AAE06000.

XX New zinc finger proteins, useful for modulating or regulating gene

XX expression and metabolic pathways in plants, e.g. for treating in the

XX plant cells a disorder that is associated with abnormal expression of the

XX target gene.

XX Example 4; Page 138-139; 156pp; English.

XX The patent discloses methods and compositions to modulate the expression

XX of a target gene in plant cells. The method involves providing plant

XX cells with a zinc finger protein (ZFP) which is capable of specifically

XX binding to a target nucleotide sequence or its complementary strand

XX within a target gene and allowing the ZFP binding to the target

XX nucleotide sequence, where the expression of the target gene in the plant

XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful

XX for modulating or regulating gene expression and metabolic pathways in

XX plants. The ZFP, fusion proteins and methods are useful in plant and

XX agricultural technology. The method is useful particularly for treating a

XX disorder in the plant cells, where the disorder is associated with

XX abnormal expression of the target gene. The present DNA sequence is the

XX partial sequence of pMal-m1 and ZFPm1 DNA

XX Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;

SQ

Query Match

Best Local Similarity 97.2%; Score 3207.2; DB 4; Length 3300;

Matches 3242; Conservative 98.2%; Pred. No. 0;

Mismatches 58; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTCGCGGTATGCGCATGATAGCCCGGAAGAGA 60

Db 1 ||||| CCGACCATCGAATGGTGC AAAA CCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACTGATATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACTGATATACGATGTGCGAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGCAAAACAGTCTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Qy 301 AGCGCCGTGCGAAATTTGTGCGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGCG 360
Db 301 AGCGCCGTGCGAAATTTGTGCGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGTCTCGATGCTAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCAATC 420
Db 361 TGGTGTCTCGATGCTAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCTGGATGACACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCTGGATGACACAGGATGCCA 480
Qy 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGCGCTTATTTCTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGCGCTTATTTCTGATGCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTTTCTCCATGAGACGCTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCATGAGACGCTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGTCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGCATTTGGTCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCCATGTCGGTTCACAAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCCATGTCGGTTCACAAACCAATGCAATGCTGA 780
Qy 781 ATGAGGGCATCGTCCCACTGGATGCTGGTTGCCAACGATCGATGGCGTGGGCGCAA 840
Db 781 ATGAGGGCATCGTCCCACTGGATGCTGGTTGCCAACGATCGATGGCGTGGGCGCAA 840
Qy 841 TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db 841 TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Qy 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAATCAAAAGGATTTTC 960
Db 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGCACTGGCAGCAGCAGGTTT 1140
Db |||||

Db 1081 CGCAACCGCCTCTCCCGCGCGTGGCGGATTCATTAATGCACTGGCAGCAGCAGGTTT 1140
Qy 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCGTGCAACGATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCGTGCAACGATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCATTCGGAAGCTGTGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 1320
Db 1261 TCAGGACGCATTCGGAAGCTGTGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 1320
Qy 1321 TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTTCGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTTCGCGGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACCAATTTTCAACAGGAAACAGCCAGTCCGTTTTAGTGTTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACCAATTTTCAACAGGAAACAGCCAGTCCGTTTTAGTGTTTTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTAATGA AAACTGAAGAGGTAAACTGCTAATCTCG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTAATGA AAACTGAAGAGGTAAACTGCTAATCTCG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGGCAAAAGCGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGGCAAAAGCGTTCCAGGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTCGGATGCGGTACGTACAAACGCGAGCTGATTGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTCGGATGCGGTACGTACAAACGCGAGCTGATTGCTTACCGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACGTAATTCACCTGGCGCTGATGCTGCTGACGCGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACGTAATTCACCTGGCGCTGATGCTGCTGACGCGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGATTAAGAACGCTGGGCGTGGATTAACGCTGGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGATTAAGAACGCTGGGCGTGGATTAACGCTGGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGGTTGACCTGATTAAGAAACAAACACATGAATGCAAGACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGGTTGACCTGATTAAGAAACAAACACATGAATGCAAGACCGATTAC 2160
Qy 2161 TCCATCGAGAACGTCCTTTTAATAAGGCGGAAACGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGAGAACGTCCTTTTAATAAGGCGGAAACGCGATGACCATCAACGGCCCGTGG 2220

```
QY 2221 GCATGGTCAACATCGACACCAAGAGTGAATTATGTGTAAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2221 GCATGGTCAACATCGACACCAAGAGTGAATTATGTGTAAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2281 AAGGTCACACCATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Db |||||
QY 2281 AAGGTCACACCATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Db |||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACGTGAGAGTCTG 2400
Db |||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACGTGAGAGTCTG 2400
Db |||||
QY 2401 GAAGCGGTAAATAAAGACAAACCGCTGGTGGTGGCTGAGCTGAAGTCTTACGAGGAAG 2460
Db |||||
QY 2401 GAAGCGGTAAATAAAGACAAACCGCTGGTGGTGGCTGAGCTGAAGTCTTACGAGGAAG 2460
Db |||||
QY 2461 TTGGCGAAGATCCACGATTTGCGGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db |||||
QY 2461 TTGGCGAAGATCCACGATTTGCGGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db |||||
QY 2581 GCCAGCGTGTGATGATGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db |||||
QY 2581 GCCAGCGTGTGATGATGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db |||||
QY 2641 AACACACACAAATACAAATAACAACCTCCGGATCGAGGAGGATTTTCAGATTC 2700
Db |||||
QY 2641 AACACACACAAATACAAATAACAACCTCCGGATCGAGGAGGATTTTCAGATTC 2700
Db |||||
QY 2701 GGATCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db |||||
QY 2701 GGATCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db |||||
QY 2761 CCGGAATGTGTAAAGTCTTCTCAGGATGCTGGCCACCTGTGTTCCGCCACCGAGTACCAC 2820
Db |||||
QY 2761 CCGGAATGTGTAAAGTCTTCTCAGGATGCTGGCCACCTGTGTTCCGCCACCGAGTACCAC 2820
Db |||||
QY 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAACTTTTAGCACCGCGGTCC 2880
Db |||||
QY 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAACTTTTAGCACCGGTCCAGAAC 2880
Db |||||
QY 2881 CTGTGCGCCATCAACGACTCATATCTGGGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db |||||
QY 2941 AAGTCTTTTCAGCCAGAGCTCCAGCTGTGGTGGCCCAACCAAGTACTACACCGGGGAGAAG 3000
Db |||||
QY 2941 AAGTCTTTTCAGCCAGAGCTCCAGCTGTGGTGGCCCAACCAAGTACTACACCGGGGAGAAG 3000
Db |||||
QY 3001 CCCTATGCTTGTCCGGGAATGTGTAAGTCTTTCAGCCAGAGAGCTCCCTGTGGTGGCCAC 3060
Db |||||
QY 3001 CCCTATGCTTGTCCGGGAATGTGTAAGTCTTTCAGCCAGAGAGTACCTGTGGTGGCCAC 3060
Db |||||
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGT 3120
Db |||||
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGT 3120
Db |||||
QY 3121 GACTGCGCGACCTTGTTCGCCATCAACGACTCATATCTGGCGAGAGGCAATACAAATGT 3180
Db |||||
QY 3121 GACTGCGCGACCTTGTTCGCCATCAACGACTCATATCTGGCGAGAGGCAATACAAATGT 3180
Db |||||
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTCGTCCGGCACCAACGACTCAC 3240
Db |||||
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCGGTCTGACAACTCTCGTCCGGCACCAACGACTCAC 3240
Db |||||
QY 3241 ACCGGTAAAAAACTAGTGGCGAGCGCGGCGAGTACCCGTACGAGTTCGGGACTAGCT 3300
Db |||||
QY 3241 ACCGGTAAAAAACTAGTGGCGAGCGCGGCGAGTACCCGTACGAGTTCGGGACTAGCT 3300
Db |||||
```

```
RESULT 4
RAD11589
ID AAD11589 standard; DNA; 3300 BP.
XX
AC AAD11589;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m2 and ZFPm2 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
CDS 2719..3270
FT /*tag= a
FT /product= "ZFPm2 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"
XX
DN WO200152620-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US001817.
XX
PR 21-JAN-2000; 2000US-0177468P.
PR 21-JUL-2000; 2000US-00620897.
XX
(SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI Barbas CF, Stege JT, Guan X, Dalmia B;
XX
```

DR WPI; 2001-465325/50.
XX P-PSDB; AAE06002.
PT New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 140-142; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m2 and ZFPm2 DNA
XX
SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;

Query Match 97.2%; Score 3207.2; DB 4; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGATGGTGCAGAAACCTTTCGGGTATGGCATGATAGGCCCGGAAGAGA 60
Db 1 CCGACACCATCGATGGTGCAGAAACCTTTCGGGTATGGCATGATAGGCCCGGAAGAGA 60

Qy 61 GTCGAATTCAGGGTGTGAATGTGAACACAGTAACGTATACGATCTCGCAGAGTATGCCG 120
Db 61 GTCGAATTCAGGGTGTGAATGTGAACACAGTAACGTATACGATCTCGCAGAGTATGCCG 120

Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTCGAAAA 180

Qy 181 CGCGGGAAAAAGTGAAGCGGGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db 181 CGCGGGAAAAAGTGAAGCGGGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240

Qy 241 AACAACTGCGGGCAAAACAGTGTCTGATTTGGGCTTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGCAAAACAGTGTCTGATTTGGGCTTGGCCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGGCCGTGCGCAAAATTTGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGGCCGTGCGCAAAATTTGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360

Qy 361 TGGTGGTGTGATGTAGAGCAAGCGGGTTCGAAGCGCTGTAAGCGGGTGCACAATC 420
Db 361 TGGTGGTGTGATGTAGAGCAAGCGGGTTCGAAGCGCTGTAAGCGGGTGCACAATC 420

Qy 421 TTCTCGCCCAACGGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCCCAACGGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480

Qy 481 TTGCTGTGGAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540

Qy 541 CACCCATCAACAGTATATTTCTCCCATGAAGACGGTACCGCATGCGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATATTTCTCCCATGAAGACGGTACCGCATGCGCGTGGAGCATC 600

Qy 601 TGGTGGCATGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db 601 TGGTGGCATGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660

Qy 661 CGCCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
Db 661 CGCCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720

Qy 721 CGGAACGGGAAGGAGCTGGAGTCCCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGAGCTGGAGTCCCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780

Qy 781 ATGAGGGCATCGTTCCTCCACTCGCATGCTGGTGGCAACGATCAGATGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCTCCACTCGCATGCTGGTGGCAACGATCAGATGCGCTGGGCGCAA 840

Qy 841 TGCGGCCATTAACGAGTCCGGGCTGCGGTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGCGGCCATTAACGAGTCCGGGCTGCGGTGGTGGCGATATCTCGGTAGTGGGATACG 900

Qy 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACGAGTTTC 960
Db 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACGAGTTTC 960

Qy 961 GCCTGCTGGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Db 961 GCCTGCTGGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA 1020

Qy 1021 AGGCAATCAGCTGTGTCCTCTCACTGGTGAAGAAAAACCAACCTTGGGCGCCAAATA 1080
Db 1021 AGGCAATCAGCTGTGTCCTCTCACTGGTGAAGAAAAACCAACCTTGGGCGCCAAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATCATTAATCGAGTGGCAGCAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATCATTAATCGAGTGGCAGCAGCAGGTTT 1140

Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200

Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGATGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGATGCAAGTGCACCAATGCTTCTGGCG 1260

Qy 1261 TCAGGCAGCATCGGAAGCTGTGGTGGCTGTCAGGTCGTAATCACTCGCATTAATTCG 1320
Db 1261 TCAGGCAGCATCGGAAGCTGTGGTGGCTGTCAGGTCGTAATCACTCGCATTAATTCG 1320

Qy 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCGACATATAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCGACATATAACGGTT 1380

Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440

Qy 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAACACCGAGTCCGTTTAGGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAACACCGAGTCCGTTTAGGTTTTCACGA 1500

Qy 1501 GCATTTCAACCAAGGACCATAGATTAATGAAAGAGGTAAGGTAATCTGGTAACTCGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATTAATGAAAGAGGTAAGGTAATCTGGTAACTCGG 1560

Qy 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680

Qy 1681 CGGCACTGGCGATGGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGTGTGGCTAC 1740
Db 1681 CGGCACTGGCGATGGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGTGTGGCTAC 1740

```
QY 1741 GCTCAATCTGGCTCTTGGCTGAATATCAACCCCGGACAAAGCGTTCACAGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTCTTGGCTGAATATCAACCCCGGACAAAGCGTTCACAGCAAGCTGTAT 1800
QY 1801 CCGTTTACTCGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACTCGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCCGCAAAACCTTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCCGCAAAACCTTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGAGAAAGTGAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGAGAAAGTGAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTGGTACCTGATTAAGAACTGAAGAACTGAAGTACGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTGGTACCTGATTAAGAACTGAAGAACTGAAGTACGACACCGGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGCAACACAGCATGACCATCAAGCGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGCAACACAGCATGACCATCAAGCGGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGCTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGCTGTAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGCTCAACCATCAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCGCCGAGT 2340
Db 2281 AAGGCTCAACCATCAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCGCCGAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAACCGGTTAATAAGACAAACCGCTGGTGCCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAACCGGTTAATAAGACAAACCGCTGGTGCCTAGCGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGCTGCTGCTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGCTGCTGCTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGATGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGATGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAGGATTCAGAAATTC 2700
QY 2701 GGAATCTCTTCTGTCGCGGCGGCGCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGAATCTCTTCTGTCGCGGCGGCGCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGCTGTAAGTCTTTCAGGATCTCGGCGACCTGCTGCGCAACGAGCTACCCAC 2820
Db 2761 CCGGAATGCTGTAAGTCTTTCCTCTCAGAGCTCTCACCTGGTGGCCACGAGCTACCCAC 2820
QY 2821 ACGGGTGAACAAACCGTATTAATCCAGAGTGGCGCAAACTCTTTTAGCACCGCGCTCC 2880
```

```
Db 2821 ACGGGTGAACAAACCGTATTAATCCAGAGTGGCGCAAACTCTTTAGCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCATCAACGCACTCATATCTGGCGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATATCTGGCGAGAGCCATACAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTTACGACAGCTCCAGCGTGGTGGCCACCAAGTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTTCTCGTCTGCAATCTCTGTCGGCACCACGATCTCACACCGGGGAGAG 3000
QY 3001 CCTATGCTTTGTCGGGAATGTGGTAACTCTTCTCAGCCAGAGCAGCTCCCTGGTCCGCCAC 3060
Db 3001 CCTATGCTTTGTCGGGAATGTGGTAACTCTTCTCAGCCAGAGCAGTAACTCTGGTCCGCCAC 3060
QY 3061 CAGGTACCCACACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAACTCTTTTAGC 3120
Db 3061 CAGGTACCCACACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAACTCTTTTAGC 3120
QY 3121 GACTGCCCGCACCTGCTCGCCATCAACGCACTCATATCTGGCGAGAGCCATACAAATGT 3180
Db 3121 CAGCCCGCCACCTGGCCAGCCATCAACGCACTCATATCTGGCGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTTCTCCCAATCCAGCCATCTCTCGCGGACCAACAGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTTCTCTCGTCTGCAAACTCTCTCGCGGACCAACAGTACTCAC 3240
QY 3241 ACCGTTAAAAAACTAGTGGCCAGCGGCGGCGGAGTACCCGTACGAGTTCGCGACTAGCT 3300
Db 3241 ACCGTTAAAAAACTAGTGGCCAGCGGCGGCGGAGTACCCGTACGAGTTCGCGACTAGCT 3300

RESULT 5
AAD11591
ID AAD11591 standard; DNA; 3300 BP.
XX
AC AAD11591;
XX
DT 24-SBP-2001 (first entry)
XX
DE Partial sequence of pMal-m4 and ZFPm4 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
CDS 2719..3270
FT /tag= a
FT /product= "ZFPm4 protein"
FT /tranel_except= (pos:3046..3048, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
```

FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"
XX
PN WO200152620-A2.
XX
XX 26-JUL-2001.
XX 19-JAN-2001; 2001WO-US001817.
XX 21-JAN-2000; 2000US-0177468P.
XX 21-JUL-2000; 2000US-00620897.
XX (SCRI) SCRIPPS RES INST.
XX (SYGN) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CF, Stege JT, Guan X, Dalmia B;
XX WPI; 2001-465325/P.
XX DR P-PSDB; AAE06004.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX
XX Example 4; Page 145-147; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX partial sequence of pMal-m4 and 2FPm4 DNA
XX
SQ Sequence 3300 BP; 849 A; 865 C; 880 G; 706 T; 0 U; 0 Other;

Query Match 97.2%; Score 3207.2; DB 4; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 CGGACACCATCAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB |||||
DB 1 CGGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACTTATACGATTCGACAGTATGCCG 120
DB |||||
DB 61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACTTATACGATTCGACAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACCAAGTAACTTCTCGGAAAA 180
DB |||||
DB 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACCAAGTAACTTCTCGGAAAA 180

QY 181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTATCATTTCCCAACCGCGTGGCAC 240
DB |||||
DB 181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTATCATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB |||||
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCGCTCGCAAAATTTGTCGGCGCGATTAATATCTCGCGCGATCAATCGGGTGCACGC 360
DB |||||
DB 301 ACGCCGCTCGCAAAATTTGTCGGCGCGATTAATATCTCGCGCGATCAATCGGGTGCACGC 360
QY 361 TGGTGGTTCGATCGGTAGAACGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 420
DB |||||
DB 361 TGGTGGTTCGATCGGTAGAACGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATACTCCGCTGGATCCGATGACCAAGATGCCA 480
DB |||||
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATACTCCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
DB |||||
DB 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
DB |||||
DB 541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
QY 601 TGGTGGCATTTGGGTACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB |||||
DB 601 TGGTGGCATTTGGGTACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGCTGCGCTGCGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB |||||
DB 661 CGCGTCTGCGCTGCGCTGCGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGGCGACTGGAGTGCCATGTCGGTTTTCACAAACCAACATGCAAAATGCTGA 780
DB |||||
DB 721 CGGAAACGGGAAGGCGACTGGAGTGCCATGTCGGTTTTCACAAACCAACATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
DB |||||
DB 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
QY 841 TGCGCGCATTTACCGAGTCCCGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATACG 900
DB |||||
DB 841 TGCGCGCATTTACCGAGTCCCGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGGTTAAACCAACATCAAAACAGGATTTTC 960
DB |||||
DB 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGGTTAAACCAACATCAAAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAACACGAGCTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB |||||
DB 961 GCCTGTGGGGCAACACGAGCTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
DB |||||
DB 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCATTAAATGACGTGGCAGCAGAGTTT 1140
DB |||||
DB 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCATTAAATGACGTGGCAGCAGAGTTT 1140
QY 1141 CCGAGTCGAAAGCGGCGAGTGAGGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
DB |||||
DB 1141 CCGAGTCGAAAGCGGCGAGTGAGGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
QY 1201 GCACAATTTCTATGTTTGACAGCTTATCATGACGTGACGGTGCACCAATGCTTCTGGCG 1260
DB |||||
DB 1201 GCACAATTTCTATGTTTGACAGCTTATCATGACGTGACGGTGCACCAATGCTTCTGGCG 1260

QY 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGTCATTAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGTCATTAATTCG 1320
QY 1321 TGTCCGCTCAAGCGGCACTCCCGCTTCGGAATAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTCCGCTCAAGCGGCACTCCCGCTTCGGAATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATATTCTCAAAATGAGCTGTGCACATTAATCATCGGCTGTAATGTTGGA 1440
Db 1381 CTGGCAAAATATTCTCAAAATGAGCTGTGCACATTAATCATCGGCTGTAATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
QY 1501 GCACCTTCCACACAGGACCATAGATTATGAAAATCTGAAGAGGTAAATCTGGTAATCTGG 1560
Db 1501 GCACCTTCCACACAGGACCATAGATTATGAAAATCTGAAGAGGTAAATCTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATAAACTGGAAGAGAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATAAACTGGAAGAGAAATTCACACAGGTT 1680
QY 1681 CGGCAACTGCGGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCACGACAACTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCACGACAACTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTAAACCGCAAGCTGATGTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTAAACCGCAAGCTGATGTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGGACCGCCGACAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGGACCGCCGACAAACCTGGGAA 1920
QY 1921 GAGATCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAACAAACACATGAATGAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAACAAACACATGAATGAGACACCGGATTAC 2160
QY 2161 TCCATCCGAGAGCTGCTTTTAAATGAAGGCGAAACAGCGATGACCATCAACGGCCGTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTTAAATGAAGGCGAAACAGCGATGACCATCAACGGCCGTGG 2220
QY 2221 GCATGGTCCACATCCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGACCTTC 2280
Db 2221 GCATGGTCCACATCCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGACCTTC 2280
QY 2281 AAGGTTCAACCATCCAAACCGTTGCTGGTGGCTGTCAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTGCTGGTGGCTGTCAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400

Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTATTAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTATTAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAACGATTTGCGGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAACGATTTGCGGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAAACATCCGCGAGATGTCGGTCTTCTGCTGATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAAACATCCGCGAGATGTCGGTCTTCTGCTGATGCGGTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATTAACAACACTCGGATCGGATCGAGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATTAACAACACTCGGATCGGATCGAGGAAGGATTTTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGGTAAAGTCTTTCAGCGATCTCGTGGCCACTGCTGTCGCCACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAAGTCTTTCAGCGAGCAGCTCTCCTGCTGCGCCACAGCGTACCCAC 2820
QY 2821 ACCGGTGAACAAACCGTATTAATGCGCAGAGTGGCGCAAAATCTTTTAGCACACCGGCTCC 2880
Db 2821 ACCGGTGAACAAACCGTATTAATGCGCAGAGTGGCGCAAAATCTTTTAGCACACCGGCTCC 2880
QY 2881 CTGGTGGCCCATCAACCGCACTCATACTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACCGCACTCATACTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTGACGACAGCTCCAGCTGGTGGCGCCACCAACGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTGACGTATGTCGTGATCTTTCGAGGACCAACGTACTCACACCGGGGAGAAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTTGAAAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTTGAAAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
QY 3121 GACTGCGCGACCTTCTCGCCATCAACGCACTCATACTGCGGAGAGGACATCAAAATGT 3180
Db 3121 CGCAGCGATAACCTGTCGCGCCATCAACGCACTCATACTGCGGAGAGGACATCAAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTCTCCCAATCCAGCACTCTCGTCCGCGACCAAGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTTCACTTCAGGCCAATTTGTCGTCACCAAGTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCGCAGCGCCGAGTACCCGTACGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCGCAGCGCCGAGTACCCGTACGAGTTCGGGACTACGCT 3300

RESULT 6

AAH25681
ID AAH25681 standard; DNA; 8101 BP.

XX AAH25681;

XX AC

XX 05-SEP-2001 (first entry)

XX DT

XX DE Nucleotide sequence of human Cdc25C phosphatase DNA.

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
KW ss.
XX
XX Homo sapiens.
XX WO200144467-A2.
XX 21-JUN-2001.
XX
XX 13-DEC-2000; 2000WO-FR003496.
XX
XX 14-DEC-1999; 99FR-00015722.
XX 30-MAY-2000; 2000FR-00006883.
XX 21-SEP-2000; 2000FR-00012008.
XX (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
XX Goubin-Gramatica F, Ducommun B, Prevost G;
XX WPI; 2001-398152/42.
XX
XX New fusion protein of maltose-binding protein and Cdc phosphatase, useful
XX for identifying phosphatase modulators for regulating the cell cycle.
XX
XX Claim 7; Page 9-13; 56pp; French.
XX
XX The present sequence encodes human Cdc25C phosphatase, and was expressed
XX in Escherichia coli. It is used to make fusion proteins with Escherichia
XX coli maltose binding protein (MBP). The specification describes fusion
XX proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
XX are involved in activation of cyclin-dependent kinases (CDK) that control
XX cell division. The fusion proteins are used to identify modulators of the
XX specified human Cdc phosphatases, potentially useful for regulation of
XX cell division. They are also for studying physiological or
XX physiopathological activities of Cdc phosphatase
XX
XX Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;
XX
XX
XX Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCGGAAGAGA 60
XX
XX 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCGGAAGAGA 60
XX
XX 61 GTCATTCAGGGTGTGAATGTGAACCAAGTAACGTTATACGATGTCGAGAGTATGCCG 120
XX
XX 61 GTCATTCAGGGTGTGAATGTGAACCAAGTAACGTTATACGATGTCGAGAGTATGCCG 120
XX
XX 121 GTGTCCTTTATCAGACCGGTTTCCCGGTGGTGAACAGGCCAGCCAGCTTTCTGCGAANA 180
XX
XX 121 GTGTCCTTTATCAGACCGGTTTCCCGGTGGTGAACAGGCCAGCCAGCTTTCTGCGAANA 180
XX
XX 181 CCGCGGAAAGAGTGAAGCGCGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
XX
XX 181 CCGCGGAAAGAGTGAAGCGCGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
XX
XX 241 AACAACTGGCGGCAAAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
XX
XX 241 AACAACTGGCGGCAAAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
XX
XX 301 ACGCGCGTCCGAATTTGCGGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGG 360
XX
XX 301 ACGCGCGTCCGAATTTGCGGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGG 360
XX
XX 361 TGGTGTGTGCGATGTAGAACGAAGCGCGTGAAGCGCTGTAAAGCGGCGGTGCACAATC 420
XX
XX 361 TGGTGTGTGCGATGTAGAACGAAGCGCGTGAAGCGCTGTAAAGCGGCGGTGCACAATC 420
XX
XX 421 TTCTCGCGCAACGGCTCAGTGGGTGATCATTAACCTATCCGCTGGATGACCAAGGATGCCA 480

Db
421 TTCTCGCGCAACGGCTCAGTGGGTGATCATTAATCTATCCCTGGATACCAGGATGCCA 480
Qy
481 TTGCTGTGGAAGCTCGCTGCACTAATGTTCCGGGGTTATTTCTTGATCTCTGTACACAGA 540
Db
481 TTGCTGTGGAAGCTCGCTGCACTAATGTTCCGGGGTTATTTCTTGATCTCTGTACACAGA 540
Qy
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Qy
601 TGGTCGCAATGGGTGTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db
601 TGGTCGCAATGGGTGTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Qy
661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db
661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy
721 CGGAAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCAATGCAATGCTGA 780
Db
721 CGGAAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCAATGCAATGCTGA 780
Qy
781 ATGAGGGCATCGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGGGCTGGGCGCA 840
Db
781 ATGAGGGCATCGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGGGCTGGGCGCA 840
Qy
841 TGGCGCCATTAACAGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db
841 TGGCGCCATTAACAGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Qy
901 ACGATACCGAAGACAGCTCATGTTATATPCCGCGGTAAACCAACCAATCAAAAGATTTTC 960
Db
901 ACGATACCGAAGACAGCTCATGTTATATPCCGCGGTAAACCAACCAATCAAAAGATTTTC 960
Qy
961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCTGCAACTCTCTCAGGCGCAGGCGTGA 1020
Db
961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCTGCAACTCTCTCAGGCGCAGGCGTGA 1020
Qy
1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGTTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
Db
1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGTTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
Qy
1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAAATGCGAGCTGCGACACAGGTTT 1140
Db
1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAAATGCGAGCTGCGACACAGGTTT 1140
Qy
1141 CCGCACTGGAAGCGGCGAGTGAGGCGAACCGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db
1141 CCGCACTGGAAGCGGCGAGTGAGGCGAACCGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Qy
1201 GCACAAATTCATGTTTTCAGACGCTTATCATGACGCTGACCGTGCAACCAATGCTTCTGGCG 1260
Db
1201 GCACAAATTCATGTTTTCAGACGCTTATCATGACGCTGACCGTGCAACCAATGCTTCTGGCG 1260
Qy
1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTTCGTAATCACTGCAATATTCG 1320
Db
1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTTCGTAATCACTGCAATATTCG 1320
Qy
1321 TGTGCTCAAGCGGCACTCCCGCTTCTGGAATATGTTTTTGGCGCGACATCAATACGGTT 1380
Db
1321 TGTGCTCAAGCGGCACTCCCGCTTCTGGAATATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy
1381 CTGGCAAAATTTCTGAAATGAGCTGTTCACAATTAATCATCGGCTCGTATATGTTGGA 1440
Db
1381 CTGGCAAAATTTCTGAAATGAGCTGTTCACAATTAATCATCGGCTCGTATATGTTGGA 1440
Qy
1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAAGA 1500
Db
1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAAGA 1500
Qy
1501 GCATTTCACAAACGAGCCATAGATTATGAAAACTGAAGAGGTAATACTGGTAACTTCGG 1560

Db 1501 GCACCTTCACCAACAAGGACCATTAGATTATGAAAAATCGAAGAGGTAAATCTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCGCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCGCCACAGGTT 1680
QY 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACCAACGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACCAACGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTGCGAACCCGCCAAAACCTGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTGCGAACCCGCCAAAACCTGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACCTGGCGCTGGATACGCTGGCGGCAAGGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACCTGGCGCTGGATACGCTGGCGGCAAGGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACATCCAAACCGTTGCTGGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGGTCAACATCCAAACCGTTGCTGGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAGGCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAGGCTG 2400
QY 2401 GAAGCGGTTAATAAGCAAAACCGCTGGGTGCGGTAGCTGTAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGCAAAACCGCTGGGTGCGGTAGCTGTAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAGATCCACGTTATGCGGCACCATGGAACCCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATGCGGCACCATGGAACCCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACAATCCCGAGATGTCGGCTTTCTGATATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACAATCCCGAGATGTCGGCTTTCTGATATGCGGTGCTGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGCTGAGACTGCTGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGAGACTGCTGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG 2640

QY 2641 AACCAACACACATTAACAATAACAACACCTCGGGATCGAGGAAGGATTTCAAGATTC 2700
Db 2641 AACCAACACACATTAACAATAACAACACCTCGGGATCGAGGAAGGATTTCAAGATTC 2700
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 7
ABL53238
ID ABL53238 standard; DNA; 8101 BP.
XX ABL53238;
AC ABL53238;
XX
DT 17-JUN-2002 (first entry)
XX
DE Nucleotide sequence of pMAL-Hs Cdc25C.
XX
KW Cdc25C phosphatase; amine; Cdc25 phosphatase; spontaneous alopecia;
KW alopecia; proliferative disease; parasitic disease; viral infection;
KW neurodegenerative disease; myopathy; angiogenesis; psoriasis; restenosis;
KW vitamin K; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key
CDS Location/Qualifiers
1528..2677
/*tag= b
/*note= "encodes maltose binding protein (MBP)"
FT CDS 2713..4134
/*tag= a
FT /*note= "ORF of human Cdc25C"
XX
PN FR2812198-A1.
XX
PD 01-FEB-2002.
XX
PF 28-JUL-2000; 2000FR-00009900.
XX
PR 28-JUL-2000; 2000FR-00009900.
XX
PA (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
PI Prevost G, Brezak PMC, Galcera CMO, Thureau C, Goubin GF;
PI Ducommun B;
XX
DR WPI; 2002-166066/22.
XX
PT cdc25 Phosphatase inhibiting medicaments, useful e.g. for treating tumor
PT diseases, viral infections, neurodegenerative disease or alopecia,
PT containing new or known aryl-substituted amines.
XX
PS Disclosure; Page 18-22; 39pp; French.
XX
CC The present sequence represents pMAL-HsCdc25C, a plasmid encoding human
CC Cdc25C phosphatase. The specification describes the use of aryl-
CC substituted secondary or tertiary amines for the production of
CC medicaments for inhibiting Cdc25 phosphatases, especially Cdc25C
CC phosphatases. The amines of the invention are used for treating
CC spontaneous alopecia or alopecia induced by exogenous products or
CC radiation. They are also used for treating tumoral or non-tumoral
CC proliferative diseases, parasitic diseases, viral infections,
CC neurodegenerative diseases or myopathy. Examples of non-tumoral
CC proliferative diseases are e.g. angiogenesis, psoriasis or restenosis.
CC The amines are also useful for inhibiting the proliferation of
CC microorganisms (especially yeasts) and in the treatment of all disorders
CC conventionally treated with vitamin K or its derivatives
XX
SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match									
Best Local Similarity 82.0%; Score 2705.8; DB 6; Length 8101;									
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	1	CGGACACCATCAATGGTGC	AAACCTTTCCGGGTATG	GCATGATAGCCCGGAAGAGA	60				
DB	1	CGGACACCATCAATGGTGC	AAACCTTTCCGGGTATG	GCATGATAGCCCGGAAGAGA	60				
QY	61	GTCAATTCAGGTGTGAATG	TGAACCCAGTAACTG	TATACGATGTCGAGAGTATGCG	120				
DB	61	GTCAATTCAGGTGTGAATG	TGAACCCAGTAACTG	TATACGATGTCGAGAGTATGCG	120				
QY	121	GTGTCCTTATCAGACCGT	TTCCCGGTGTCGAA	CCAGCCAGCCACGCTTCTCGGAAA	180				
DB	121	GTGTCCTTATCAGACCGT	TTCCCGGTGTCGAA	CCAGCCAGCCACGCTTCTCGGAAA	180				
QY	181	CCGCGGAAAAAGTGGAA	GGCGGATGGCGAGCT	GAAATACATTCCTCAACCCGCTGGCAC	240				
DB	181	CCGCGGAAAAAGTGGAA	GGCGGATGGCGAGCT	GAAATACATTCCTCAACCCGCTGGCAC	240				
QY	241	AAACAATGCGGGCAAA	CAGTCTGTTGCTGAT	TGGCGTTGCCACCTCCAGTCTGGCCCTGC	300				
DB	241	AAACAATGCGGGCAAA	CAGTCTGTTGCTGAT	TGGCGTTGCCACCTCCAGTCTGGCCCTGC	300				
QY	301	ACGCGCGTCCGCAAA	TGTCGGCGGATTA	ATCTCGCGCCGATCAA	CTGGGTGCCAGG	360			
DB	301	ACGCGCGTCCGCAAA	TGTCGGCGGATTA	ATCTCGCGCCGATCAA	CTGGGTGCCAGG	360			
QY	361	TGGTGTGTCGATGTTAG	AAACGAGCGGCTG	AAAGCGGCGTGCA	CAATC	420			
DB	361	TGGTGTGTCGATGTTAG	AAACGAGCGGCTG	AAAGCGGCGTGCA	CAATC	420			
QY	421	TTCTCGCGCAACGCGT	CAGTGGGCTGAT	CAATTAACCTATCCGCTGGATGAC	CCAGGATGCCA	480			
DB	421	TTCTCGCGCAACGCGT	CAGTGGGCTGAT	CAATTAACCTATCCGCTGGATGAC	CCAGGATGCCA	480			
QY	481	TTGCTGTGGAAGCTGC	TGCTGACATATG	TTCCGGCGTTATTTCTGATGTTCTG	ACACGA	540			
DB	481	TTGCTGTGGAAGCTGC	TGCTGACATATG	TTCCGGCGTTATTTCTGATGTTCTG	ACACGA	540			
QY	541	CACCCATCAACAGTAT	ATTTTCTCCATGA	AGCGGTACGCGACTGGGCGTGGAGCATC	600				
DB	541	CACCCATCAACAGTAT	ATTTTCTCCATGA	AGCGGTACGCGACTGGGCGTGGAGCATC	600				
QY	601	TGGTCGCATTTGGT	CACGCAAAATCG	CGCTGTTAGCGGSCCATTTAAGTCTGCTCGG	660				
DB	601	TGGTCGCATTTGGT	CACGCAAAATCG	CGCTGTTAGCGGSCCATTTAAGTCTGCTCGG	660				
QY	661	CGCGTCTGCGTCTGG	CTGGCTGGCATAA	ATATCTCACTCGCAATCAAATTCAGCCGATAG	720				
DB	661	CGCGTCTGCGTCTGG	CTGGCTGGCATAA	ATATCTCACTCGCAATCAAATTCAGCCGATAG	720				
QY	721	CGGACGGGAAGCGGAC	TGAGTGCCATG	TCGCGTTTTCAACAAACCATGCAATGCTGA	780				
DB	721	CGGACGGGAAGCGGAC	TGAGTGCCATG	TCGCGTTTTCAACAAACCATGCAATGCTGA	780				
QY	781	ATGAGGGCATCTGTT	CCCACTGGGATG	CTGTTGCCAACGATCAGATGGCGCTGGGCGCAA	840				
DB	781	ATGAGGGCATCTGTT	CCCACTGGGATG	CTGTTGCCAACGATCAGATGGCGCTGGGCGCAA	840				
QY	841	TGCGGCCATACCGAG	TCCGGCTGCGGTT	GTGGGATATCTCGGTAGTGGGATAG	900				
DB	841	TGCGGCCATACCGAG	TCCGGCTGCGGTT	GTGGGATATCTCGGTAGTGGGATAG	900				
QY	901	ACGATACCAAGACAG	CTCATGTTATAT	CCCGCGTTTACCCCATCAACAGGATTTTC	960				
DB	901	ACGATACCAAGACAG	CTCATGTTATAT	CCCGCGTTTACCCCATCAACAGGATTTTC	960				
QY	961	GCCTGTGGGGCAAA	CCAGCGTGGAC	CGCTGTTGTCGAACTCTCTCAGGCGCCAGGGGTGA	1020				
DB	961	GCCTGTGGGGCAAA	CCAGCGTGGAC	CGCTGTTGTCGAACTCTCTCAGGCGCCAGGGGTGA	1020				
QY	1021	AGGGCAATCAGTGTT	GGCCCGTCTCA	CTCGTGGAAAGAAAAACCCCTGGGCGCCCAATA	1080				

DB	1021	AGGGCAATCAGTGTTGG	CCCGTCTCACTGGT	GAAAGAAAAACCCCTGGGCGCCCAATA	1080				
QY	1081	CGCAAAACCGCCTCT	CCCCCGCGTGGCCG	ATTCATTAAATGCAGCTGGCACGACAGGTTT	1140				
DB	1081	CGCAAAACCGCCTCT	CCCCCGCGTGGCCG	ATTCATTAAATGCAGCTGGCACGACAGGTTT	1140				
QY	1141	CCGACTGGAAGCGG	CAGTGAGGCAACG	CAATTAATGTGAGTTAGTCTACTCATTTAG	1200				
DB	1141	CCGACTGGAAGCGG	CAGTGAGGCAACG	CAATTAATGTGAGTTAGTCTACTCATTTAG	1200				
QY	1201	GCACAATTTCTATG	TTTGACAGCTTAT	CATCGACTGACGCTGCACCAATGTTCTTGCGG	1260				
DB	1201	GCACAATTTCTATG	TTTGACAGCTTAT	CATCGACTGACGCTGCACCAATGTTCTTGCGG	1260				
QY	1261	TCAGGACGCCAT	TCGGAAGCTGTG	GTATGGCTGTGCAGGTCTGTAAATCACTGCATAAATTCG	1320				
DB	1261	TCAGGACGCCAT	TCGGAAGCTGTG	GTATGGCTGTGCAGGTCTGTAAATCACTGCATAAATTCG	1320				
QY	1321	TGTGCTCAAGCG	CACCTCCCGTTCT	CGATTAATGTTTTTTGGCGCGACATCATNAACGGTT	1380				
DB	1321	TGTGCTCAAGCG	CACCTCCCGTTCT	CGATTAATGTTTTTTGGCGCGACATCATNAACGGTT	1380				
QY	1381	CTGGCAAAATAT	TCTGAATGAGCT	GTTCACAATTAATCATCGGCTCGTATATGTTGTCGA	1440				
DB	1381	CTGGCAAAATAT	TCTGAATGAGCT	GTTCACAATTAATCATCGGCTCGTATATGTTGTCGA	1440				
QY	1441	ATTGTGAGCGG	ATAACAAATTT	CACACAGGAAACGCCAGTCCGTTTTAGGTGTTTTACGA	1500				
DB	1441	ATTGTGAGCGG	ATAACAAATTT	CACACAGGAAACGCCAGTCCGTTTTAGGTGTTTTACGA	1500				
QY	1501	GCACCTTCAAC	AAAGGACCAT	AGATTAAGAAAATGAAAGAGTAAACTGGTAATCTCG	1560				
DB	1501	GCACCTTCAAC	AAAGGACCAT	AGATTAAGAAAATGAAAGAGTAAACTGGTAATCTCG	1560				
QY	1561	ATTAAACGGG	ATAAGGCTATA	ACGCTCCGCTGAAGTCGGTAAGAAATTCAGAAAGAT	1620				
DB	1561	ATTAAACGGG	ATAAGGCTATA	ACGCTCCGCTGAAGTCGGTAAGAAATTCAGAAAGAT	1620				
QY	1621	ACCGGAAATTA	AGTCAACCGTT	GAGCATCCGGATAAATCGGAAGAGAAAATTC	1680				
DB	1621	ACCGGAAATTA	AGTCAACCGTT	GAGCATCCGGATAAATCGGAAGAGAAAATTC	1680				
QY	1681	GGGCAACTGGCG	ATGGCCCTGAC	ATTATCTTCTGGGCACACGACGCTTTGGTGGCTAC	1740				
DB	1681	GGGCAACTGGCG	ATGGCCCTGAC	ATTATCTTCTGGGCACACGACGCTTTGGTGGCTAC	1740				
QY	1741	GCTCAATCTGCG	CTGTTGGCTGAA	ATCAACCGGACAAAGCGTTCAGGACAAAGCTGTAT	1800				
DB	1741	GCTCAATCTGCG	CTGTTGGCTGAA	ATCAACCGGACAAAGCGTTCAGGACAAAGCTGTAT	1800				
QY	1801	CCGTTTACCTGG	ATGCCGTAC	TTTCAACCGGCAAGCTGATTCCTTACCCGATCGCTGT	1860				
DB	1801	CCGTTTACCTGG	ATGCCGTAC	TTTCAACCGGCAAGCTGATTCCTTACCCGATCGCTGT	1860				
QY	1861	GAAAGCGTTAT	CGCTGATTTAA	CAAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920				
DB	1861	GAAAGCGTTAT	CGCTGATTTAA	CAAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920				
QY	1921	GAGATCCCGCG	CTGGATTAAG	AACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980				
DB	1921	GAGATCCCGCG	CTGGATTAAG	AACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980				
QY	1981	CTGCAAGAAAC	CGTACTCT	CACTGGCGCTGATTTGCTGCTAGCGGGGTATTGCGTTCAAG	2040				
DB	1981	CTGCAAGAAAC	CGTACTCT	CACTGGCGCTGATTTGCTGCTAGCGGGGTATTGCGTTCAAG	2040				
QY	2041	TATGAAAAAC	CGCAAGTAC	CAATTAAGACGTTGGCGCTGGATTAACGCTGGCGCGAAAGCG	2100				
DB	2041	TATGAAAAAC	CGCAAGTAC	CAATTAAGACGTTGGCGCTGGATTAACGCTGGCGCGAAAGCG	2100				
QY	2101	GCTCTGACCT	TCCTGTTGAC	CTGATTAAGAAACAAACATGAATGCAGACACCGATTAC	2160				

```
Db 2101 GGTCTGACCTTCTGTGTTGACCTGATTAAAAACAACACATGAATGACGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCTTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACGACGAAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACGACGAAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCTGTGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGTGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTATGCCGCCACCATGGAAACGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTATGCCGCCACCATGGAAACGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAATATCCGCGAGATGTCGCTTTCTGTTGATGCGGTGCTACTGCGGTGATCAAGGCC 2580
Db 2521 CCGAATATCCGCGAGATGTCGCTTTCTGTTGATGCGGTGCTACTGCGGTGATCAAGGCC 2580
QY 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCAGTCTG 2640
Db 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCAGTCTG 2640
QY 2641 AACAAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCCCTCT 2709
Db 2701 GGATCCCTCT 2709

RESULT 8
AD023608
ID AD023608 standard; DNA; 7259 BP.
AC
XX
AC AD023608;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-ToxoP30del18 (82-294aa) fusion protein.
XX
KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
FN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PP 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
```

PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;

XX WPI; 2004-304563/28.

DR P-PSDB; ADO23609.

XX

PT Novel purified polypeptide having sequence identity to amino acid
sequence of P30 antigen of *Toxoplasma gondii* e.g. Toxo30del13C,
Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of Igm or Igg
antibodies to *Toxoplasma gondii*.

XX Example 2; Fig 22; 114pp; English.

PS The invention relates to a purified P30 antigen (I) chosen from 3 fully
defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
amino acids added to the C-terminus of the amino acid sequence of
Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
which at least one of the five C-terminal cysteine amino acids of the
amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
with alanine, or comprising the amino acid sequence chosen from MBP-
Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
fusion proteins. (I) is useful for detecting the presence of Igm
antibodies to *Toxoplasma gondii* in a test sample, which involves
contacting the test sample suspected of containing the Igm antibodies
with a composition comprising (I) and detecting the presence of (I)/Igm
antibody complexes. The present sequence represents DNA encoding a MBP-
ToxoP30 fusion protein of the invention.

XX Sequence 7259 BP; 1870 A; 1822 C; 1926 G; 1641 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCACACCATCGATGGTGCAGAAACCTTTCCGGGTATGGATGATAGCGCCCGAAGAGA 60

Db 1 CCACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGATGATAGCGCCCGAAGAGA 60

QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACTGATACGATGTCGAGAGTATGCCG 120

Db 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACTGATACGATGTCGAGAGTATGCCG 120

QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGCCAGCCACGTTTCTGCGAAAA 180

Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGCCAGCCACGTTTCTGCGAAAA 180

QY 181 CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTAATCTCCACCCGCGTGGCAC 240

Db 181 CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTAATCTCCACCCGCGTGGCAC 240

QY 241 AACAACTGGCGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

Db 241 AACAACTGGCGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

QY 301 AGCGCGCGTGCAGAAATTTGCGGGCGGATTAATCTCGCGCCGATCACTGGGTGCCAGCG 360

Db 301 AGCGCGCGTGCAGAAATTTGCGGGCGGATTAATCTCGCGCCGATCACTGGGTGCCAGCG 360

QY 361 TGGTGTGTCGATGGTAGAACGAAGCGGCTCGAAGCCCTGTAAGCGGCGGTGCACATC 420

Db 361 TGGTGTGTCGATGGTAGAACGAAGCGGCTCGAAGCCCTGTAAGCGGCGGTGCACATC 420

QY 421 TTCTCGCGCAACCGGTGAGTCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480

Db 421 TTCTCGCGCAACCGGTGAGTCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480

QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCTTATTTCTGATGTCTTGACCAGA 540

Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCTTATTTCTGATGTCTTGACCAGA 540

QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

Db 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

QY 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
DB 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTCTGCGTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGGTCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGCATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGCATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCGATCGTCCCACTCGCATGCTGGTGGCCAAACGATCAGATGGCGTGGCGGCA 840
DB 781 ATGAGGCGATCGTCCCACTCGCATGCTGGTGGCCAAACGATCAGATGGCGTGGCGGCA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATAG 900
DB 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACGACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTCAGTTAGTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTCAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCACTGCAAGTGCAGGTCACCAATGCTCTGGCG 1260
DB 1201 GCACAATCTCATGTTTGACAGCTTATCATCACTGCAAGTGCAGGTCACCAATGCTCTGGCG 1260
QY 1261 TCAGGCGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATAAATTCG 1320
DB 1261 TCAGGCGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGGTT 1380
DB 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCACTCGGCTCGTATAATGTGGA 1440
DB 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCACTCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTCAAGA 1500
DB 1441 ATTGTGAGCGGATAACAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTCAAGA 1500
QY 1501 GCATTTACCAACAGGACATAGATTAATGAAAACTGAAGAGGTAAACTGTTAATCTGG 1560
DB 1501 GCATTTACCAACAGGACATAGATTAATGAAAACTGAAGAGGTAAACTGTTAATCTGG 1560
QY 1561 ATTTACGGCGATAAAGGCTATACCGTCTCGCTGAGTTCGTAAGAAATTCAGAAAGAT 1620
DB 1561 ATTTACGGCGATAAAGGCTATACCGTCTCGCTGAGTTCGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTTAAAGTCAACCGTTCGAGTCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTTAAAGTCAACCGTTCGAGTCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGGCCCTGACATTAATCTCTGGGGCAACGACGCGTTTGGTGGCTAC 1740

RESULT 9
ADO23613
ID ADO23613 standard; DNA; 7322 BP.
XX AC ADO23613;

DB 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGAACAAGCGTTTCCAGGAACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGAACAAGCGTTTCCAGGAACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAACAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTATTAACAACAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACCGTGGCGTGGATAACGCTGGCGCGAAGCG 2100
DB 2041 TATGAAACCGCAAGTACGACATTAAGACCGTGGCGTGGATAACGCTGGCGCGAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAACCAACACATGAATTCAGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAACCAACACATGAATTCAGACACCGATTAC 2160
QY 2161 TCCATCGAGAAGTGCCTTTAATAAGCGAACAAGCGATGACCATCAACGGCCCGTGG 2220
DB 2161 TCCATCGAGAAGTGCCTTTAATAAGCGAACAAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACCAAGCAAGTGAATTAATGCTGTAACCGGTACTGCGACCTTC 2280
DB 2221 GCATGCTCAACATCGACACCAAGCAAGTGAATTAATGCTGTAACCGGTACTGCGACCTTC 2280
QY 2281 AAGGTCACACCATCAAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
DB 2281 AAGGTCACACCATCAAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGAAGAAGTCTG 2400
DB 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG 2460
DB 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG 2460
QY 2461 TTGGCAAGAGATCCACGTTATTGCGCCCACTGGAACCAACGCCAGAAAGGTGAAATCATG 2520
DB 2461 TTGGCAAGAGATCCACGTTATTGCGCCCACTGGAACCAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACAATCCCGCAGATGTCGCTTTTGGTATGCGGTGCTGCTGCTGATCAACCGCC 2580
DB 2521 CCGAACAATCCCGCAGATGTCGCTTTTGGTATGCGGTGCTGCTGCTGATCAACCGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTTAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAACAATAACAACCAACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700
DB 2641 AACCAACAACAATAACAATAACAACCAACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700

XX 01-JUL-2004 (first entry)
 XX DNA encoding MBP-Toxop30del110 (52-284aa) fusion protein.
 XX DE
 XX P30 antigen; Toxop30del13C; Toxop30del2C; Toxop30 MIX1;
 KW MBP-Toxop30del13C (52-300aa); MBP-Toxop30del14C (52-294aa); MBP-Toxop30MIX1;
 KW Toxoplasma gondii; de; gene.
 XX OS
 XX Toxoplasma gondii.
 OS Synthetic.
 XX US2004067239-A1.
 XX 08-APR-2004.
 PD
 XX 02-OCT-2002; 2002US-00263153.
 PF
 XX 02-OCT-2002; 2002US-00263153.
 PR
 XX (MAIN/) MAINE G T.
 PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 XX
 PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX WPI; 2004-304563/28.
 DR P-PSDB; ADO23614.
 DR
 XX Novel purified polypeptide having sequence identity to amino acid
 PT sequence of p30 antigen of Toxoplasma gondii e.g. Toxop30del13C,
 PT Toxop30del2C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX
 PS Example 2; Fig 25; 114pp; English.
 XX
 CC The invention relates to a purified p30 antigen (I) chosen from 3 fully
 CC defined Toxop30del13C, Toxop30del2C and Toxop30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxop30del13C p30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxop30del13C p30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxop30del13C (52-300aa), MBP-Toxop30del14C (52-294aa) and MBP-Toxop30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC Toxop30 fusion protein of the invention.
 XX
 SQ Sequence 7322 BP; 1887 A; 1844 C; 1934 G; 1657 T; 0 U; 0 Other;
 Query Match 81.5%; Score 2690.4; DB 12; Length 7322;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGTTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
 DB
 QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACTTATACGATGTCGACAGATATGCCG 120
 DB 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACTTATACGATGTCGACAGATATGCCG 120
 QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGTTTCTGCCGAAA 180
 DB 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGTTTCTGCCGAAA 180
 QY 181 CGCGGGAAAGTGGAGCGCGCATGCGGAGCTGAATTACATTCCTCCACCGCGTGGCAC 240
 DB 181 CGCGGGAAAGTGGAGCGCGCATGCGGAGCTGAATTACATTCCTCCACCGCGTGGCAC 240

QY 241 AACAACTGGCGGCAAAACAGTCGTGCTGATGCGGTGCGACCTCCAGTCTGCGCCCTGC 300
 DB AACAACTGGCGGCAAAACAGTCGTGCTGATGCGGTGCGACCTCCAGTCTGCGCCCTGC 300
 QY 301 ACGCGCGTCGCAAAATTTGTGCGCGCATTAATCTCGCGCGCATCAACTGGGTGCGACGC 360
 DB ACGCGCGTCGCAAAATTTGTGCGCGCATTAATCTCGCGCGCATCAACTGGGTGCGACGC 360
 QY 361 TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCGCTGTAAAGCGCGCGTGCACAATC 420
 DB TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCGCTGTAAAGCGCGCGTGCACAATC 420
 QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCACTCCGCTGGATGACAGATGCCA 480
 DB TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCACTCCGCTGGATGACAGATGCCA 480
 QY 481 TTGCTGTGGAAGTGCCTGCACCTAAATGTTCCGGCGTTATTTCTTTGATGTTCTCTGACCAGA 540
 DB TTGCTGTGGAAGTGCCTGCACCTAAATGTTCCGGCGTTATTTCTTTGATGTTCTCTGACCAGA 540
 QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGCTGGGCGTGGAGCATC 600
 DB CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGCTGGGCGTGGAGCATC 600
 QY 601 TGGTCGCATTTGGGTACCCAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTCTGCTCGG 660
 DB TGGTCGCATTTGGGTACCCAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTCTGCTCGG 660
 QY 661 CGCGTCTGCGTCTGGGTGGTGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
 DB CGCGTCTGCGTCTGGGTGGTGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
 QY 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
 DB CGGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
 QY 781 ATGAGGCGCATCGTTCCCACTCGCATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
 DB ATGAGGCGCATCGTTCCCACTCGCATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
 QY 841 TGGCGCGCATTAACGAGTCGCGGCTCGCGGTGGTGGTGGATATCTCGGTAGTGGGATACG 900
 DB TGGCGCGCATTAACGAGTCGCGGCTCGCGGTGGTGGTGGATATCTCGGTAGTGGGATACG 900
 QY 901 ACATACCGAAGACACTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
 DB ACATACCGAAGACACTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
 QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGCGTGA 1020
 DB GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGCGTGA 1020
 QY 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACCCACCTGGCGCCCAATA 1080
 DB AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACCCACCTGGCGCCCAATA 1080
 QY 1081 CGCAAAACCGCCTCTCCCCCGCGTTGGCGCATTCATTAATGAGCTGGCACACAGGTTT 1140
 DB CGCAAAACCGCCTCTCCCCCGCGTTGGCGCATTCATTAATGAGCTGGCACACAGGTTT 1140
 QY 1141 CCGGACTGGAAGCGGCGCATGAGCGCAACGCAATTAATGTAAGTACTCACTCATTAG 1200
 DB CCGGACTGGAAGCGGCGCATGAGCGCAACGCAATTAATGTAAGTACTCACTCATTAG 1200
 QY 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCGCAATGCTTCTGGCG 1260
 DB GCACAATTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCGCAATGCTTCTGGCG 1260
 QY 1261 TCAGGACGACCATCGGAAGCTGTGGTATGCGTGTGAGGTTCGTAATCACTGTCATTAATTCG 1320
 DB TCAGGACGACCATCGGAAGCTGTGGTATGCGTGTGAGGTTCGTAATCACTGTCATTAATTCG 1320

Qy	1321	TGTCGCTCAAGGGCGCACTCCCGTTCTGGATAAATGTTTTTTTGGCGCGACATCATAAACGGTT	1380
Db	1321	TGTCGCTCAAGGGCGCACTCCCGTTCTGGATAAATGTTTTTTTGGCGCGACATCATAAACGGTT	1380
Qy	1381	CTGSCAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATAATGTTGGGA	1440
Db	1381	CTGSCAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATAATGTTGGGA	1440
Qy	1441	ATTGTGAGCGGATAACAAATTTTCACACAGGAAAACAGCCAGTCCGTTTAGTGTGTTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTTCACACAGGAAAACAGCCAGTCCGTTTAGTGTGTTTTTCACGA	1500
Qy	1501	GCATTTACCAACAGAGACATAGATTATGAAAACCTGAAAAGTAACTCGTAACTCTGG	1560
Db	1501	GCATTTACCAACAGAGACATAGATTATGAAAACCTGAAAAGTAACTCGTAACTCTGG	1560
Qy	1561	ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGGAATTAAGTACCGTTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCCAACAGGTT	1680
Db	1621	ACCGGAATTAAGTACCGTTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCCAACAGGTT	1680
Qy	1681	CGGCAACTCGCGATGCGCCCTGCACATTTATCTTTGGGCAACAGCCGTTTGGTGGCTAC	1740
Db	1681	CGGCAACTCGCGATGCGCCCTGCACATTTATCTTTGGGCAACAGCCGTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTCGGCTGTTGGCTGTAATCACCCCGACAAAGCGTTTCCAGGACAAGCTGTAT	1800
Db	1741	GCTCAATCTCGGCTGTTGGCTGTAATCACCCCGACAAAGCGTTTCCAGGACAAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTACAAACGGCAAGCTGATTTGCTTACCCEGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCGGTACGTTACAAACGGCAAGCTGATTTGCTTACCCEGATCGCTGTT	1860
Qy	1861	GAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
Qy	1921	GAGATCCGGCGCTGGATAAGACTCAAGCGAAGAGTAAAGCGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCGGCGCTGGATAAGACTCAAGCGAAGAGTAAAGCGCGCTGATGTTTCAAC	1980
Qy	1981	CTGCAAGAAACCGTACTTCACTGCGCGCTGATTTGCTCTGACGGGGGTTATGGTTCAAG	2040
Db	1981	CTGCAAGAAACCGTACTTCACTGCGCGCTGATTTGCTCTGACGGGGGTTATGGTTCAAG	2040
Qy	2041	TATGAAAACCGGCAAGTACGACATTAAGAAGTGGGCGTGGATTAACGCTGGCGGAAGCG	2100
Db	2041	TATGAAAACCGGCAAGTACGACATTAAGAAGTGGGCGTGGATTAACGCTGGCGGAAGCG	2100
Qy	2101	GGTCTGACCTTCTGGTTGACCTGATTAACAAACACATGATGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGGTTGACCTGATTAACAAACACATGATGAGACACCGATTAC	2160
Qy	2161	TCCATCGCAAGTCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAAGTCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Qy	2221	GCATGGTCCAAACATCGACACCGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGGTCCAAACATCGACACCGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGGTCAACATCGAAACCGTTCTGGCGTGTGAGCGCGAGGTATTAACCGCGCCAGT	2340
Db	2281	AAGGGTCAACATCGAAACCGTTCTGGCGTGTGAGCGCGAGGTATTAACCGCGCCAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGGTTCTCGAAAACCTATCTGCTGACTGATGAGGTCG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGGTTCTCGAAAACCTATCTGCTGACTGATGAGGTCG	2400
Qy	2401	GAAGCGGTTATAAAGACAAACCGCTGGGTGCGCTAGCGTTGAAGTCTTACGAGGAAGAG	2460

Db	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGGCGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAAGATCCACGTATTTCGGCGCACCATGGAATAAAGCGCCAGAAAGGTGAAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTATTTCGGCGCACCATGGAATAAAGCGCCAGAAAGGTGAAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGCTGAGACTGTCGATGAAGCCCTCAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGCTGAGACTGTCGATGAAGCCCTCAAGACGCGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAACAACAACAATAACCAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2641	AACAACAACAACAATAACCAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
RESULT 10			
ADO23603			
ID	ADO23603	standard; DNA; 7352 BP.	
XX	AC	ADO23603;	
XX	DT	01-JUL-2004 (first entry)	
XX	XX	DNA encoding MBP-ToxoP30del14C(52-294aa) fusion protein.	
XX	DE		
XX	KW	P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;	
XX	KW	MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;	
XX	KW	Toxoplasma gondii; ds; gene.	
OS	OS	Toxoplasma gondii.	
OS	OS	Synthetic.	
XX	FN	US2004067239-A1.	
XX	XX		
XX	PD	08-APR-2004.	
XX	XX		
XX	PF	02-OCT-2002; 2002US-00263153.	
XX	PR	02-OCT-2002; 2002US-00263153.	
XX	XX	(MAIN/) MAINE G T.	
PA	PA	(PATE/) PATEL C B.	
PA	PA	(GINS/) GINSBURG S R.	
PA	PA	(BLIE/) BLIESE T R.	
XX	XX		
PI	PI	Maine GT, Patel CB, Ginsburg SR, Bliese TR;	
XX	XX		
DR	DR	WPI; 2004-304563/28.	
DR	DR	P-PSDB; ADO23604.	
XX	PT		
XX	PT	Novel purified polypeptide having sequence identity to amino acid	
PT	PT	sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,	
PT	PT	Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG	
PT	PT	antibodies to Toxoplasma gondii.	
XX	XX		
PS	PS	Example 2; Fig 19; 114pp; English.	
XX	XX		
CC	CC	The invention relates to a purified p30 antigen (I) chosen from 3 fully	
CC	CC	defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6	
CC	CC	amino acids added to the C-terminus of the amino acid sequence of	
CC	CC	Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in	
CC	CC	which at least one of the five C-terminal cysteine amino acids of the	
CC	CC	amino acid sequence of Toxo30del13C P30 antigen sequence is substituted	
CC	CC	with alanine, or comprising the amino acid sequence chosen from MBP-	
CC	CC	Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1	
CC	CC	fusion proteins. (I) is useful for detecting the presence of IgM	
CC	CC	antibodies to Toxoplasma gondii in a test sample, which involves	
CC	CC	contacting the test sample suspected of containing the IgM antibodies	

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
XX Toxop30 fusion protein of the invention.

SQ Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

Query Match	81.5%;	Score	2690.4;	DB	12;	Length	7352;
Best Local Similarity	99.8%;	Pred. No.	0;				
Matches	2694;	Conservative	0;	Mismatches	6;	Indels	0;
Gaps	0;						

QY	1	CGGACCAATCAATGGTGCAGAAACCTTTGCGGGTATGCGATGATGCGCCGGAAGAGA	60
DB	1	CGGACCAATCAATGGTGCAGAAACCTTTGCGGGTATGCGATGATGCGCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCGG	120
DB	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTACGTTATACGATGTCGACAGTATGCGG	120
QY	121	GTGTCCTTATCAGACCGGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAA	180
DB	121	GTGTCCTTATCAGACCGGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAA	180
QY	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATTAATCCCAACCGGTGGCAC	240
DB	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATTAATCCCAACCGGTGGCAC	240
QY	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTCGCAATTTGTCGGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
DB	301	ACGCGCGTCGCAATTTGTCGGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
QY	361	TGGTGGTTCGATGATGAACGAGCGCGTGAAGCTGTAAGCGCGGTGACAAATC	420
DB	361	TGGTGGTTCGATGATGAACGAGCGCGTGAAGCTGTAAGCGCGGTGACAAATC	420
QY	421	TTCTCGCGCAACGGGTCACTAGTGGGTGATCAATTAATCTCCGCTGGATGACACGAGTGC	480
DB	421	TTCTCGCGCAACGGGTCACTAGTGGGTGATCAATTAATCTCCGCTGGATGACACGAGTGC	480
QY	481	TTGCTGTGGAAGCTCGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTGACACGA	540
DB	481	TTGCTGTGGAAGCTCGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTGACACGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGCGTGGAGCATC	600
QY	601	TGGTGCATTTGGGTACACAGCAATTCGGCTGTTAGCGGGCCATTAAGTTCGTCTCGG	660
DB	601	TGGTGCATTTGGGTACACAGCAATTCGGCTGTTAGCGGGCCATTAAGTTCGTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGGATGAATATCTCACTCCCAATCAATTTTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGGATGAATATCTCACTCCCAATCAATTTTCAGCCGATAG	720
QY	721	CGGAACCGGAAGGCGATCGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780
DB	721	CGGAACCGGAAGGCGATCGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCGATGTTCCCACTCGCATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA	840
DB	781	ATGAGGCGATGTTCCCACTCGCATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA	840
QY	841	TGCGCGCATTAACGAGTCCGGGCTGGCGTGGTGGCGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGCGCATTAACGAGTCCGGGCTGGCGTGGTGGCGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACGGAAGACAGTCTATGTTATATCCGCGGTTAACCAATCAACAGGATTTTC	960
DB	901	ACGATACGGAAGACAGTCTATGTTATATCCGCGGTTAACCAATCAACAGGATTTTC	960

QY	961	GCCTGCTGGGCAAAACAGGGTGCACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGCAAAACAGGGTGCACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGCAATACAGCTGTTGCCGCTCTCACTGCTGTAAGAAAGAAACACCCCTGGCGCCAAATA	1080
DB	1021	AGGCAATACAGCTGTTGCCGCTCTCACTGCTGTAAGAAAGAAACACCCCTGGCGCCAAATA	1080
QY	1081	CGCAAAACCGCTCTCTCCCGCGGCTTGCGCGATTCATTAATGACGCTGGCAGCAGAGTTT	1140
DB	1081	CGCAAAACCGCTCTCTCCCGCGGCTTGCGCGATTCATTAATGACGCTGGCAGCAGAGTTT	1140
QY	1141	CCGACCTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCAATTAG	1200
DB	1141	CCGACCTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCAATTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCCATCGGAAGCTGTGCTATGCTGTCAGGTCGTAAATCACTGCAATAATTCG	1320
DB	1261	TCAGGACGCCATCGGAAGCTGTGCTATGCTGTCAGGTCGTAAATCACTGCAATAATTCG	1320
QY	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT	1380
DB	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGCTCGTATAATGTTGGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGCTCGTATAATGTTGGGA	1440
QY	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500
QY	1501	GCATTTCAACCAACAGGACCATAGATTATGAACCTGAGAGGTTAACTGGTAACTCTCG	1560
DB	1501	GCATTTCAACCAACAGGACCATAGATTATGAACCTGAGAGGTTAACTGGTAACTCTCG	1560
QY	1561	ATTAAACGGGATTAAGGCTATTAACCGTCTCGCTGAAAGTCCGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACGGGATTAAGGCTATTAACCGTCTCGCTGAAAGTCCGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACCAAGTTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACCAAGTTT	1680
QY	1681	GGGCAACTGGCGATGGCGCTTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
DB	1681	GGGCAACTGGCGATGGCGCTTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGCAAAAGCGTTCCAGACAGCTGTAT	1800
DB	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGCAAAAGCGTTCCAGACAGCTGTAT	1800
QY	1801	CGTTTTACCTGGGATGCGGTACGTTTACACGCGCAAGCTGATTTGCTTACCGATCGCTGTT	1860
DB	1801	CGTTTTACCTGGGATGCGGTACGTTTACACGCGCAAGCTGATTTGCTTACCGATCGCTGTT	1860
QY	1861	GAGCGTTATTCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920
DB	1861	GAGCGTTATTCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAGAGGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAGAACTGAAGAGGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGGCTGATTTGCTGTCGACGGGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGGCGGCTGATTTGCTGTCGACGGGGTTATCGGTTCAAG	2040

[illegible]

PA	(BLIE/) BLIESE T R.	
XX	Maine GT, Patel CB, Ginsburg SR, Bliese TR;	
XX	WPI; 2004-304563/28.	
XX	P-PSDB; ADO23599.	
XX	Novel purified polypeptide having sequence identity to amino acid	
PT	sequence of P30 antigen of <i>Toxoplasma gondii</i> e.g. Toxo30del13C,	
PT	Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG	
PT	antibodies to <i>Toxoplasma gondii</i> .	
XX	Example 2; Fig 15; 114pp; English.	
XX	The invention relates to a purified P30 antigen (I) chosen from 3 fully	
CC	defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6	
CC	amino acids added to the C-terminus of the amino acid sequence of	
CC	Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in	
CC	which at least one of the five C-terminal cysteine amino acids of the	
CC	amino acid sequence of Toxo30del13C P30 antigen sequence is substituted	
CC	with alanine, or comprising the amino acid sequence chosen from MBP-	
CC	Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1	
CC	fusion proteins. (I) is useful for detecting the presence of IgM	
CC	antibodies to <i>Toxoplasma gondii</i> in a test sample, which involves	
CC	contacting the test sample suspected of containing the IgM antibodies	
CC	with a composition comprising (I) and detecting the presence of (I)/IgM	
CC	antibody complexes. The present sequence represents DNA encoding a MBP-	
CC	ToxoP30 fusion protein of the invention.	
XX	Sequence 7370 BP; 1901 A; 1855 C; 1949 G; 1665 T; 0 U; 0 Other;	
SQ	Query Match 81.5%; Score 2690.4; DB 12; Length 7370;	
	Best Local Similarity 99.8%; Pred. No. 0;	
	Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
Qy	1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60	
Db	1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60	
Qy	61 GTCAATTTCAGGTCGTGAATCTGTGAACCAGTAAAGCTTATACGATGTCGAGAGTATGCCG 120	
Db	61 GTCAATTTCAGGTCGTGAATCTGTGAACCAGTAAAGCTTATACGATGTCGAGAGTATGCCG 120	
Qy	121 GTGTCTCTTATCAGACCGCTTTTCGCGGTGTGTGAACGAGCCAGCCACGTTTCTGCGAAAA 180	
Db	121 GTGTCTCTTATCAGACCGCTTTTCGCGGTGTGTGAACGAGCCAGCCACGTTTCTGCGAAAA 180	
Qy	181 CGCGGGAAAAAGTGGAAAGCGCGCATGCGGAGCTGAATTACATCCCAACCGGTGGCAC 240	
Db	181 CGCGGGAAAAAGTGGAAAGCGCGCATGCGGAGCTGAATTACATCCCAACCGGTGGCAC 240	
Qy	241 AACCACTGGCGGGCAACAGTCGTTGCTGATGGGTGTTCCACCTCCAGTCTGGCCCTGC 300	
Db	241 AACCACTGGCGGGCAACAGTCGTTGCTGATGGGTGTTCCACCTCCAGTCTGGCCCTGC 300	
Qy	301 ACGGCGCGTCGCAAAATTGTGCGGCGCATTTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360	
Db	301 ACGGCGCGTCGCAAAATTGTGCGGCGCATTTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360	
Qy	361 TGTGTGTGTCGATCGGTGAAGCAAGACGCGCTCGAAGCTGTAAAGCGGCGGTGCACAATC 420	
Db	361 TGTGTGTGTCGATCGGTGAAGCAAGACGCGCTCGAAGCTGTAAAGCGGCGGTGCACAATC 420	
Qy	421 TTCTCTCGCAACCGTCTAGTGGGCTGATCAATTAATCTCCGTCGATGACAGGATGCCA 480	
Db	421 TTCTCTCGCAACCGTCTAGTGGGCTGATCAATTAATCTCCGTCGATGACAGGATGCCA 480	
Qy	481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGGTATTTCTTGATGTCTCTGACCGA 540	
Db	481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGGTATTTCTTGATGTCTCTGACCGA 540	
Qy	541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACCGCATGCGCGGTGGAGCATC 600	

Db	541	CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATC	600
Qy	601	TGGTCGCAATTGGGTACACGAAATCGCGCTGTGTAGCGGCCCAATTAAGTTCTGTCTCGG	660
Db	601	TGGTCGCAATTGGGTACACGAAATCGCGCTGTGTAGCGGCCCAATTAAGTTCTGTCTCGG	660
Qy	661	CGGCTCTGGCTGTGGCTGGCTGGCTGAATAATCTCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGGCTCTGGCTGTGGCTGGCTGGCTGAATAATCTCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAACGGGAAGCGGACGTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAACGGGAAGCGGACGTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAATGCTGA	780
Qy	781	ATGAGGGCATCTTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGAA	840
Db	781	ATGAGGGCATCTTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGAA	840
Qy	841	TGCGCGCCATTACCGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCCATTACCGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTATTATCCCGCGGTTAAACCAACATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTATTATCCCGCGGTTAAACCAACATCAACAGGATTTTC	960
Qy	961	GCCTGCTGGGGAACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGGAACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTGCGCTCTCACTGTGTAAGAAAGAAACCAACCTTGCGCGCCAAATA	1080
Db	1021	AGGGCAATCAGCTGTGCGCTCTCACTGTGTAAGAAAGAAACCAACCTTGCGCGCCAAATA	1080
Qy	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGACGTGCGACGACAGTTT	1140
Db	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGACGTGCGACGACAGTTT	1140
Qy	1141	CCGACTGGAAGCGGCGAGTGAGGCAACCGCAATTAATGTAGTTAGTCACTCATTTAG	1200
Db	1141	CCGACTGGAAGCGGCGAGTGAGGCAACCGCAATTAATGTAGTTAGTCACTCATTTAG	1200
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATGCTGACGCTGACCAATGCTCTCTGCGG	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATGCTGACGCTGACCAATGCTCTCTGCGG	1260
Qy	1261	TCAGGACGCCATCGGAAGCTGTGTATGGCTGTGAGGTGCAAGTCACTGCATAATTTCG	1320
Db	1261	TCAGGACGCCATCGGAAGCTGTGTATGGCTGTGAGGTGCAAGTCACTGCATAATTTCG	1320
Qy	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGATATGTTTTTGGCGCGACATCAACCGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGATATGTTTTTGGCGCGACATCAACCGTT	1380
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA	1440
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA	1440
Qy	1441	ATTGTGACGGATCAAAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAACGA	1500
Db	1441	ATTGTGACGGATCAAAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAACGA	1500
Qy	1501	GCATCTCACCAACAGGACCATAGATTATGAAACTGAAGAAGTAACTGGTAAATCTGG	1560
Db	1501	GCATCTCACCAACAGGACCATAGATTATGAAACTGAAGAAGTAACTGGTAAATCTGG	1560
Qy	1561	ATTAAACGGCGATTAAGGCTATAACGGTCTCGCTGAAAGTAAAGTAACTGGAAGAAGAT	1620
Db	1561	ATTAAACGGCGATTAAGGCTATAACGGTCTCGCTGAAAGTAAAGTAACTGGAAGAAGAT	1620
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTTCCACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTTCCACAGGTT	1680

RESULT 12

ADO23639

ID ADO23639 standard; DNA; 7370 BP.

Qy	1681	CGCGAACTGCGGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGCGAACTGCGGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGCGCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGCGCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCCGTACGTTTCAACGGCAAGCTGATTTGCTTACCCTGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCCGTACGTTTCAACGGCAAGCTGATTTGCTTACCCTGATCGCTGTT	1860
Qy	1861	GAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCGAAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCGAAACCCGCCAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGCTAAGACGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGCTAAGACGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACGGGGTTATGGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACGGGGTTATGGTTCAAG	2040
Qy	2041	TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATAAACGCTGGCGGAAGCG	2100
Db	2041	TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATAAACGCTGGCGGAAGCG	2100
Qy	2101	GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACATGATGACAGACCGGATTAC	2160
Db	2101	GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACATGATGACAGACCGGATTAC	2160
Qy	2161	TCCATGCGAAGACTGCTTTTAAAGGCGAAACAGCGATGACCAATCAACGCCCGTGG	2220
Db	2161	TCCATGCGAAGACTGCTTTTAAAGGCGAAACAGCGATGACCAATCAACGCCCGTGG	2220
Qy	2221	GCATGCTCAACATCGACACCGCAAGTGAATTAATGCTGTAACGGTACTGCGGACCTTC	2280
Db	2221	GCATGCTCAACATCGACACCGCAAGTGAATTAATGCTGTAACGGTACTGCGGACCTTC	2280
Qy	2281	AAGGGTCAACCATTCAAAACCGTTGCGTGTGAGCGCAGGTATTAACGCCCGCAGT	2340
Db	2281	AAGGGTCAACCATTCAAAACCGTTGCGTGTGAGCGCAGGTATTAACGCCCGCAGT	2340
Qy	2341	CCGAAACAAAGAGCTGCAAAAGAGTTCTCGAAAACATCTCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAAACAAAGAGCTGCAAAAGAGTTCTCGAAAACATCTCTGCTGACTGATGAAGTCTG	2400
Qy	2401	GAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG	2460
Db	2401	GAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAGAGTCCAGTATTCGCCACCATGGAAGAAACGCCCGAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAGAGTCCAGTATTCGCCACCATGGAAGAAACGCCCGAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCCTGTTCTGTTATGCGTGGTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCCTGTTCTGTTATGCGTGGTACTGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGTGTGACATGTGTCGATGAAGCCCTGAAAGAGCGCGACAGTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGTGTGACATGTGTCGATGAAGCCCTGAAAGAGCGCGACAGTAAATTCGAGCTCG	2640
Qy	2641	AACAAACAAACAAATAAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2641	AACAAACAAACAAATAAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700

XX AC ADO23639;
 XX DT 01-JUL-2004 (first entry)
 XX DE DNA encoding MBP-ToxoP30MIX1 fusion protein.
 XX KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
 KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; ds; gene.
 XX OS Toxoplasma gondii.
 OS Synthetic.
 XX PN US2004067239-A1.
 XX PD 08-APR-2004.
 XX PF 02-OCT-2002; 2002US-00263153.
 XX PR 02-OCT-2002; 2002US-00263153.
 XX PA (MAIN/) MAINE G T.
 PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX WPI; 2004-304563/28.
 DR P-FSDB; ADO23640.
 XX Novel purified polypeptide having sequence identity to amino acid
 PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX Example 5; Fig 31; 114pp; English.
 XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
 CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;
 SQ
 Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGGTGCAAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
 DB 1 CCGACACCATCGAATGGTGCAAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
 QY 61 CTCATTCAGGGTGTGATGTGAACCACTTATACCATCTCCAGAGTATGCG 120
 DB 61 CTCATTCAGGGTGTGATGTGAACCACTTATACCATCTCCAGAGTATGCG 120
 QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAAACAGGCCAGCCACGTTTCTCGCAAAA 180
 DB 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAAACAGGCCAGCCACGTTTCTCGCAAAA 180
 QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240

DB 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTTATTTCCCAACCGCGTGGCAC 240
 QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGGCCCTTGC 300
 DB 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGGCCCTTGC 300
 QY 301 ACGCGCGTCCCAAAATTTGTCGCGCGATTAATCTCGCGCGATCAATCTGGGTGGCACGG 360
 DB 301 ACGCGCGTCCCAAAATTTGTCGCGCGATTAATCTCGCGCGATCAATCTGGGTGGCACGG 360
 QY 361 TGGTGGTGTGATGTGAACGAGCGCGTCGAGCGCTGTAAAGCGCGGTGACACAATC 420
 DB 361 TGGTGGTGTGATGTGAACGAGCGCGTCGAGCGCTGTAAAGCGCGGTGACACAATC 420
 QY 421 TTCTCGCGCAACGGGTGATGAGTGGGTGATCAATTAATCTATCCGCTGGATGACCAATGCA 480
 DB 421 TTCTCGCGCAACGGGTGATGAGTGGGTGATCAATTAATCTATCCGCTGGATGACCAATGCA 480
 QY 481 TTGCTGTGGAAGTGCCTGCACTAATGTTCCGGGTGATTTCTTGTGATGCTCTGACCAAGA 540
 DB 481 TTGCTGTGGAAGTGCCTGCACTAATGTTCCGGGTGATTTCTTGTGATGCTCTGACCAAGA 540
 QY 541 CACCATCAACAGTATTTTCTCCCATGAAGCGGTACGCGTGGCGGTGGAGCATC 600
 DB 541 CACCATCAACAGTATTTTCTCCCATGAAGCGGTACGCGTGGCGGTGGAGCATC 600
 QY 601 TGGTGGCATTTGGGTGATGAGTGGGTGATGAGTGGGTGATGAGTGGGTGATGAGTGGGTGATG 660
 DB 601 TGGTGGCATTTGGGTGATGAGTGGGTGATGAGTGGGTGATGAGTGGGTGATGAGTGGGTGATG 660
 QY 661 CGGCTGTGCGTGTGGGTGATGAGTGGGTGATGAGTGGGTGATGAGTGGGTGATGAGTGGGTGATG 720
 DB 661 CGGCTGTGCGTGTGGGTGATGAGTGGGTGATGAGTGGGTGATGAGTGGGTGATGAGTGGGTGATG 720
 QY 721 CGGACCGGGAAGCGACTGGAGTGGCATGCGGTTTCAACAAACCATGCAAAATGCTGA 780
 DB 721 CGGACCGGGAAGCGACTGGAGTGGCATGCGGTTTCAACAAACCATGCAAAATGCTGA 780
 QY 781 ATGAGGGCATGCTTCCCACTGCGATGCTGTTGTCGCAACGATGCGCTGGCGCGCA 840
 DB 781 ATGAGGGCATGCTTCCCACTGCGATGCTGTTGTCGCAACGATGCGCTGGCGCGCA 840
 QY 841 TGGCGCGCATTTACCGAGTCCGGGTGCGGATGCTGCGGATGCTGCGGATGCTGCGGATGCTGCGG 900
 DB 841 TGGCGCGCATTTACCGAGTCCGGGTGCGGATGCTGCGGATGCTGCGGATGCTGCGGATGCTGCGG 900
 QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCATCAACAGGATTTTC 960
 DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCATCAACAGGATTTTC 960
 QY 961 GCCTGTGCGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
 DB 961 GCCTGTGCGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
 QY 1021 AGGGCAATCAGCTGTGTGCGCGTCTCACTGGTGAAGAAAAAACCCCTTGGCGCCCAATA 1080
 DB 1021 AGGGCAATCAGCTGTGTGCGCGTCTCACTGGTGAAGAAAAAACCCCTTGGCGCCCAATA 1080
 QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGCGACGACAGGTTT 1140
 DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGCGACGACAGGTTT 1140
 QY 1141 CCCGACTCGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
 DB 1141 CCCGACTCGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
 QY 1201 GCACAATTTCTCATGTTTTCAGAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGCGG 1260
 DB 1201 GCACAATTTCTCATGTTTTCAGAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGCGG 1260
 QY 1261 TCAGGACGCCATCGGAAGCTGTGATGCGTGTGAGGTGCGTAATCATCTGCATTAATTTTCG 1320

Db 1261 TCAGGAGCGCATCGGAAGCTGTGGTATAGCTGTGCAGGTCGTAATACTACTGATATTCG 1320
Qy 1321 TGTGCTCAAGGCGCACTCCGCTTCCTGGATAATGTTTTTTCGCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCGCTTCCTGGATAATGTTTTTTCGCGACATCAATAACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTGTGAACAATTAATCATCGGCTCGTATATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTGTGAACAATTAATCATCGGCTCGTATATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAACAGCCAGTCGGTTAGGTGTTTTTCAGCA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAACAGCCAGTCGGTTAGGTGTTTTTCAGCA 1500
Qy 1501 GCATTTCAACAAAGGACCATAGCATATGAAATCTGAAGAGTAAATCTGTAATCTGG 1560
Db 1501 GCATTTCAACAAAGGACCATAGCATATGAAATCTGAAGAGTAAATCTGTAATCTGG 1560
Qy 1561 ATTAACGGGATAAAGGCTATTAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATTAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGCTGGTGAATCAACCCGGAACAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGCTGGTGAATCAACCCGGAACAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACGCGAAGCTGATGTTTACCCCATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACGCGAAGCTGATGTTTACCCCATCGCTGTT 1860
Qy 1861 GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCGAACCAGGCTGATGTTTCAAC 1920
Db 1861 GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCGAACCAGGCTGATGTTTCAAC 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAGTAAAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAGTAAAGCGCGCTGATGTTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGTTCTGACGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGTTCTGACGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAGAACCGCAAGTACGATTAAGACGCTGGCGTGGATACGCTGGCGGAAGCG 2100
Db 2041 TATGAAGAACCGCAAGTACGATTAAGACGCTGGCGTGGATACGCTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCCAGAGCTGCTTTTATAAAGGCGAACAACGCGATGACCATCAACGCGCGGCTGG 2220
Db 2161 TCCATCCAGAGCTGCTTTTATAAAGGCGAACAACGCGATGACCATCAACGCGCGGCTGG 2220
Qy 2221 GCATGGTCCAAATCCGACAGCAAGTGAATTTATGGTGAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCCGACAGCAAGTGAATTTATGGTGAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGTCACATCAACAAACCGTTCGTTGGCGTCTGAGCGCAGGATTAACGCGCGCAGT 2340
Db 2281 AAGGTCACATCAACAAACCGTTCGTTGGCGTCTGAGCGCAGGATTAACGCGCGCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAGAGTTCCTCGAACAATCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAGAGTTCCTCGAACAATCTATCTGCTGACTGATGAAGTCTG 2400

RESULT 13

ADO23649

ID ADO23649 standard; DNA; 7370 BP.

XX AC ADO23649;

XX DT 01-JUL-2004 (first entry)

XX DE DNA encoding MBP-Toxop30MIX5 fusion protein.

XX KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;

XX KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;

XX KW Toxoplasma gondii; ds; gene.

XX OS Toxoplasma gondii.

XX OS Synthetic.

XX PN US2004067239-A1.

XX PD 08-APR-2004.

XX PF 02-OCT-2002; 2002US-00263153.

XX PR 02-OCT-2002; 2002US-00263153.

XX PA (MAIN/) MAINE G T.

XX PA (PATE/) PATEL C B.

XX PA (GINS/) GINSBURG S R.

XX PA (BLIE/) BLIESE T R.

XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;

XX DR WPI; 2004-304563/28.

XX DR P-PSDB; ADO23650.

XX PT Novel purified polypeptide having sequence identity to amino acid

XX PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,

XX PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG

XX PT antibodies to Toxoplasma gondii.

XX PS Example 5; Fig 35; 114pp; English.

XX CC The invention relates to a purified p30 antigen (I) chosen from 3 fully

XX CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6

XX CC amino acids added to the C-terminus of the amino acid sequence of

XX CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in

XX CC which at least one of the five C-terminal cysteine amino acids of the

XX CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted

XX CC with alanine, or comprising the amino acid sequence chosen from MBP-

XX CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1

XX CC fusion proteins. (I) is useful for detecting the presence of IgM

CC antibodies to *Toxoplasma gondii* in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (i) and detecting the presence of (i)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATGAAATGGTGCAGAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGA 60
DB 1 CCGACACCATGAAATGGTGCAGAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGA 60

QY 61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG 120
DB 61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG 120

QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG 180
DB 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG 180

QY 181 CGCGGAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
DB 181 CGCGGAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240

QY 241 AACAACTGCGGGCAAAACAGTCTGTTGCTGATTTGGGCTTGGCACCTCCAGTCTGGGCCCTGC 300
DB 241 AACAACTGCGGGCAAAACAGTCTGTTGCTGATTTGGGCTTGGCACCTCCAGTCTGGGCCCTGC 300

QY 301 AGCGCCGTCGCAAAATGTCGCGCGATTAATCTCGCGCGGATCAATTCGCGGTCGACGG 360
DB 301 AGCGCCGTCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCAATTCGCGGTCGACGG 360

QY 361 TGGTGGTGTGATGTAGAACGAAGCGGCTGGAAGCGGCTGTAAGCGGCGTGCACATC 420
DB 361 TGGTGGTGTGATGTAGAACGAAGCGGCTGGAAGCGGCTGTAAGCGGCGTGCACATC 420

QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGTATGCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGTATGCA 480

QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTTGACCCAGA 540

QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

QY 601 TGGTGCATTTGGTGCACGAAATCGCGCTGTTAGCGGCGCCATTTAAGTTCTGCTCGG 660
DB 601 TGGTGCATTTGGTGCACGAAATCGCGCTGTTAGCGGCGCCATTTAAGTTCTGCTCGG 660

QY 661 CGCGTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAACGGGAAGCGGACTGGATGCCATGTCGCGTTTTCACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGGACTGGATGCCATGTCGCGTTTTCACAAACCATGCAAAATGCTGA 780

QY 781 ATGAGGGCATCTTCCCACTGGCATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 840
DB 781 ATGAGGGCATCTTCCCACTGGCATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 840

QY 841 TGGCGCCATTAACCGAGTCCGGGCTCGGCTGGGATATCTCGTAGTGGGATACG 900
DB 841 TGGCGCCATTAACCGAGTCCGGGCTCGGCTGGGATATCTCGTAGTGGGATACG 900

QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCATCAAAACAGGATTTTC 960

DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACGAGCGTGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACGAGCGTGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020

QY 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACACACCTTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACACACCTTGGCGCCCAATA 1080

QY 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCAGACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCAGACAGGTTT 1140

QY 1141 CCCGACTGGAAGCGGCGAGTGAGGCGAAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGGCGAAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200

QY 1201 GCACAAATTCATGTTTGCACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATTCATGTTTGCACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260

QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGCAGTTCGTAATCACTGCATAAATTCG 1320
DB 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGCAGTTCGTAATCACTGCATAAATTCG 1320

QY 1321 TGTGCTCAAGCGGCACTCCCGTCTTGATTAATGTTTTTGGCGCGCAATCAATAAACGGTT 1380
DB 1321 TGTGCTCAAGCGGCACTCCCGTCTTGATTAATGTTTTTGGCGCGCAATCAATAAACGGTT 1380

QY 1381 CTGGCAATATCTGAAATGAGCTTTCACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1440
DB 1381 CTGGCAATATCTGAAATGAGCTTTCACACAGGAAATGAGCTTTCACACAGGAAATGAGTGTTCGA 1440

QY 1441 ATTGTAGCGGATTAACAAATTTTCACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
DB 1441 ATTGTAGCGGATTAACAAATTTTCACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500

QY 1501 GCACCTTCCCAACAGGACCATAGATTAATGAAAACTGGAAGAGGTAATCTGTTAATCTCGG 1560
DB 1501 GCACCTTCCCAACAGGACCATAGATTAATGAAAACTGGAAGAGGTAATCTGTTAATCTCGG 1560

QY 1561 ATTAACGGCGATTAAGGCTATAACCGTCTCGCTGAGTGGTAAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATTAAGGCTATAACCGTCTCGCTGAGTGGTAAAGAAATTCGAGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAATCTGGAAGAGAAATTCGCCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAATCTGGAAGAGAAATTCGCCACAGGTT 1680

QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740

QY 1741 GCTCAATCTGGGCTGTTGGCTGAAATCAACCCCGGCAAAAGCGTTCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGGCTGTTGGCTGAAATCAACCCCGGCAAAAGCGTTCAGGACAAGCTGTAT 1800

QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860

QY 1861 GAAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
DB 1861 GAAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920

QY 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980

QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGAGCGGGGTTATCGGTTCAAG 2040

```
Db 1981 CTGAAGAACCGTACTTACCTGGCGCTGATTGCTGCTGAGCGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAGACGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGAAGCG 2100
Db 2041 TATGAAGACGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACAAACACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACAAACACATGAATGCGAGACACCGATTAC 2160
Qy 2161 TCCATCGAAGAGCTGCTTTAATAAGCGGAACAGCGATGACATCAACAGCGCCGTGG 2220
Db 2161 TCCATCGAAGAGCTGCTTTAATAAGCGGAACAGCGATGACATCAACAGCGCCGTGG 2220
Qy 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGGTTAAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGGTTAAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGCGGAAAGATCCACGTATTGCGGCCACCATGGAACAGCCCAAGAAAGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTATTGCGGCCACCATGGAACAGCCCAAGAAAGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGGCTTCTGGTATCCGTGCTAGCTGGGTGATCAAGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGGCTTCTGGTATCCGTGCTAGCTGGGTGATCAAGCC 2580
Qy 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTG 2640
Db 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTG 2640
Qy 2641 AACAAACAAACAAATACATTAACAAACAACTCGGGATCGAGGAAGGATTCAGAAATC 2700
Db 2641 AACAAACAAACAAATACATTAACAAACAACTCGGGATCGAGGAAGGATTCAGAAATC 2700
```

RESULT 14

ADO23644 standard; DNA; 7370 BP.

XX ADO23644;

AC ADO23644;

DT 01-JUL-2004 (first entry)

DE DNA encoding MBP-Toxop30MX1X3 fusion protein.

XX P30 antigen; Toxo30del13C; Toxo30del12C; Toxop30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.

XX Toxoplasma gondii.

OS Synthetic.

XX US2004067239-A1.

XX 08-APR-2004.

XX 02-OCT-2002; 2002US-00263153.

XX 02-OCT-2002; 2002US-00263153.

XX (MAIN/) MAINE G T.

PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.

XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
XX P-PSDB; ADO23645.

XX Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.

XX Example 5; Fig 33; 114pp; English.

XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and Toxop30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.

XX Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Qy Query Match 81.5%; Score 2690.4; DB 12; Length 7370;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCACACCATCGAATGGTGCACAAACCTTTCCGGTATGGCATATAGCGCCCGGAGAGA 60

Db 1 CCACACCATCGAATGGTGCACAAACCTTTCCGGTATGGCATATAGCGCCCGGAGAGA 60

Qy 61 GTCATTCAGGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120

Db 61 GTCATTCAGGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120

Qy 121 GTCTCTTATCAGACGGTTTCCGGGTGTGAACAGGCCAGCCAGCTTTCTGCGGAAAA 180

Db 121 GTCTCTTATCAGACGGTTTCCGGGTGTGAACAGGCCAGCCAGCTTTCTGCGGAAAA 180

Qy 181 CCGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240

Db 181 CCGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGCAAAACAGTGTCTGATTGGCGTGTGCCACCTCCAGTCTGGCCCTGC 300

Db 241 AACAACTGGCGGCAAAACAGTGTCTGATTGGCGTGTGCCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGCCCGCTGCAAAATGTGCGGCGGATTAATCTCGCGCGGATCACTGGGTGCCAGCG 360

Db 301 ACGCCCGCTGCAAAATGTGCGGCGGATTAATCTCGCGCGGATCACTGGGTGCCAGCG 360

Qy 361 TGGTGTGTGTCGATGGTAGAACGAAAGCGCGTCAAGCCCTGTAAGCGCGCGTGCACAATC 420

Db 361 TGGTGTGTGTCGATGGTAGAACGAAAGCGCGTCAAGCCCTGTAAGCGCGCGTGCACAATC 420

Qy 421 TTCTCCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGTGGATGACAGAGTGCCA 480

Db 421 TTCTCCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGTGGATGACAGAGTGCCA 480

Qy 481 TTGCTGTGAAGCTGCTGCACCTAATGTTCCGCGGTTATTTCTTGATGCTCTGACGAGA 540

Db 481 TTGCTGTGAAGCTGCTGCACCTAATGTTCCGCGGTTATTTCTTGATGCTCTGACGAGA 540

QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGCGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGCGCTGGAGCATC 600
QY 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
Db 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGGCTGGCTGGCGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTGGCGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACATGAGTGCATTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACATGAGTGCATTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCACCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCGTTCACCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGCGAA 840
QY 841 TGGCGGCCAATTAACGAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCAATTAACGAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACGAAATTAATGTGAGTTAGTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACGAAATTAATGTGAGTTAGTCACTCATTAG 1200
QY 1201 GCACAATTCATGTTTGAACGCTTATCATGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGAACGCTTATCATGACTGCACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCATCGGAAGCTGTGGTATGGCTGTGAGGTTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGTGGTATGGCTGTGAGGTTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCGTTCTGGATATGTTTTTGGCGCGACATCAACGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCGTTCTGGATATGTTTTTGGCGCGACATCAACGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATACAATTTCAACAGGAACAGCGAGTCCGTTAGGTGTTTTCAAGA 1500
Db 1441 ATTGTAGCGGATACAATTTCAACAGGAACAGCGAGTCCGTTAGGTGTTTTCAAGA 1500
QY 1501 GCACCTCCACCAAGGACCATAGCATATGAAAAATCGAAGAGGTTAACTGGTAATCTGG 1560
Db 1501 GCACCTCCACCAAGGACCATAGCATATGAAAAATCGAAGAGGTTAACTGGTAATCTGG 1560
QY 1561 ATTTAAGCGGATATAAGGCTTAAACCGTCTCGCTGAAAGTTCGGAAGGATTTTCAAG 1620
Db 1561 ATTTAAGCGGATATAAGGCTTAAACCGTCTCGCTGAAAGTTCGGAAGGATTTTCAAG 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTTCCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 GCGGCAACTGCGGATGGCCCTGCACATTATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGCGGATGGCCCTGCACATTATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGGAATCAACCCGGACAAGCGTTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGGAATCAACCCGGACAAGCGTTTCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGCGATGCGTACGTTTACAACGCGAAGCTGATTGCTTACCAGATCCCTGTT 1860
Db 1801 CCGTTTACCTGCGATGCGTACGTTTACAACGCGAAGCTGATTGCTTACCAGATCCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACGCTACTTCACTGCGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCACTGCGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAAGACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAAGACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACTGATTAATAAACAACATGATGATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACTGATTAATAAACAACATGATGATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGCAACAGCGATCACCATTCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGCAACAGCGATCACCATTCAACGGCCCGTGG 2220
QY 2221 GCATGCTCCAAACATCGACACCAAGCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCCAAACATCGACACCAAGCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCAAACATCAAAACCGTTCTGGTGGCTGAGCGCAGGATTTAAACGCCGCCAGT 2340
Db 2281 AAGGTCAAACATCAAAACCGTTCTGGTGGCTGAGCGCAGGATTTAAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTTGCGGCCACCATGGAAGAAACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTTGCGGCCACCATGGAAGAAACGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGTCAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700

AD023594
 ID ADO23594 standard; DNA; 7403 BP.
 XX AC ADO23594;
 XX DT 01-JUL-2004 (first entry)
 XX DE DNA encoding MBP-ToxoP30del2(52-311aa) fusion protein.
 XX KW P30 antigen; Toxo30del13C; Toxo30del2C; ToxoP30 MIX1;
 XX KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 XX KW Toxoplasma gondii; db; gene.
 XX OS Toxoplasma gondii.
 XX OS Synthetic.
 XX PN US2004067239-A1.
 XX PD 08-APR-2004.
 XX PF 02-OCT-2002; 2002US-00263153.
 XX PR 02-OCT-2002; 2002US-00263153.
 XX PA (MAIN/) MAINE G T.
 XX PA (PATE/) PATEL C B.
 XX PA (GINS/) GINSBURG S R.
 XX PA (BLIE/) BLIESE T R.
 XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX DR WPI; 2004-304563/28.
 XX DR P-PSDB; ADO23595.
 XX PT Novel purified polypeptide having sequence identity to amino acid
 PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX Example 2; Fig 11; 114pp; English.
 XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
 CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX Sequence 7403 BP; 1907 A; 1863 C; 1961 G; 1672 T; 0 U; 0 Other;
 SQ
 Query Match 81.5%; Score 2690.4; DB 12; Length 7403;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGTTGCAAAACCTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
 DB 1 CCGACACCATCGAATGTTGCAAAACCTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
 QY 61 GTCAATTACAGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
 DB 61 GTCAATTACAGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
 QY 121 GTGTCCTTATCAGACCGTTTCCGGGTGTGAACACAGGTATACGATGTGCGAGAGTATGCCG 180
 DB 121 GTGTCCTTATCAGACCGTTTCCGGGTGTGAACACAGGTATACGATGTGCGAGAGTATGCCG 180

181 CGCGGGAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCCCAACCGCGTGGCAC 240
 DB 181 CGCGGGAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCCCAACCGCGTGGCAC 240
 QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTGC 300
 DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTGC 300
 QY 301 ACGCGCGTGCAGAAATTTGTCCGCGGATTAATCTCCGCGCGATCAACTGGGTGCCACGC 360
 DB 301 ACGCGCGTGCAGAAATTTGTCCGCGGATTAATCTCCGCGCGATCAACTGGGTGCCACGC 360
 QY 361 TGGTGGTGTGATGGTAGAAGCGCGCTCGAAGCCTGTAAGCGCGCGGTGCACAATC 420
 DB 361 TGGTGGTGTGATGGTAGAAGCGCGCTCGAAGCCTGTAAGCGCGCGGTGCACAATC 420
 QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGATGCCA 480
 DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGATGCCA 480
 QY 481 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGCGTTAATTTCTTGATGTTCTTGACGAGA 540
 DB 481 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGCGTTAATTTCTTGATGTTCTTGACGAGA 540
 QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTAGCGACTGGGCGTGGAGGATC 600
 DB 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTAGCGACTGGGCGTGGAGGATC 600
 QY 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGTGTAGCGGCCCATTAAGTTCTGTTCTCGG 660
 DB 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGTGTAGCGGCCCATTAAGTTCTGTTCTCGG 660
 QY 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 DB 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 QY 721 CGGAAACGGGAAGCGAGCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
 DB 721 CGGAAACGGGAAGCGAGCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
 QY 781 ATGAGGCGATCGTTCCACCTGGCATCTGTTGCCAACGATCAGATGGCGTGGGCGCAA 840
 DB 781 ATGAGGCGATCGTTCCACCTGGCATCTGTTGCCAACGATCAGATGGCGTGGGCGCAA 840
 QY 841 TGGCGCCATTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
 DB 841 TGGCGCCATTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
 QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
 DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
 QY 961 GCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCGGATATCTCAGGGCCAGGGGTGA 1020
 DB 961 GCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCGGATATCTCAGGGCCAGGGGTGA 1020
 QY 1021 AGGGCAATCAGCTGTTGGCCGCTCACTGGTCAAAAGAAAACCAACCTGGGCGCCCAATA 1080
 DB 1021 AGGGCAATCAGCTGTTGGCCGCTCACTGGTCAAAAGAAAACCAACCTGGGCGCCCAATA 1080
 QY 1081 CGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGGTGGCAGCAGAGTTT 1140
 DB 1081 CGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGGTGGCAGCAGAGTTT 1140
 QY 1141 CCGGATCGAAACGGGCGATGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
 DB 1141 CCGGATCGAAACGGGCGATGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
 QY 1201 GCACAATTTCTATGTTGACAGCTTATCATCGACTGCACGGTGCACCAATGTTCTGGCG 1260
 DB 1201 GCACAATTTCTATGTTGACAGCTTATCATCGACTGCACGGTGCACCAATGTTCTGGCG 1260

```
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTCATATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTCATATTCG 1320
QY 1321 TGTGCTCAAGCGCACATCCCGTTCCTGATATATGTTTTTGGCGGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACATCCCGTTCCTGATATATGTTTTTGGCGGACATCATAAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTTCACATTAATCATCGGCTCGTATATATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTTCACATTAATCATCGGCTCGTATATATGTTGGA 1440
QY 1441 ATTGTAGCGGATTAACAATTTTACACAGGAACACCGAGTCCGTTTAGGTGTTTTTCAGA 1500
Db 1441 ATTGTAGCGGATTAACAATTTTACACAGGAACACCGAGTCCGTTTAGGTGTTTTTCAGA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATATATGAAACTGAAAGAGGTAAACTCGTAACTCGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATATATGAAACTGAAAGAGGTAAACTCGTAACTCGG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGCGGATGGCCCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGCGGATGGCCCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGACAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGACAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTCCTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTCCTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTCGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAARACGGCAAGTACGACATTAAGACGTGGCGGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAARACGGCAAGTACGACATTAAGACGTGGCGGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGAACATCAACGGCCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGAACATCAACGGCCCCGTGG 2220
QY 2221 GCATGCTCCACATCGACACCGCAAGTGAATATGTTGTTAAGCGTACTGCCGACCTTC 2280
Db 2221 GCATGCTCCACATCGACACCGCAAGTGAATATGTTGTTAAGCGTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTTCGTTGGCGGTGCTGAGCGCAGGTATTAAACGCCCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTCGTTGGCGGTGCTGAGCGCAGGTATTAAACGCCCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
```

```
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGTGCCGTAGCGCTGAAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGTGCCGTAGCGCTGAAAGTCTTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGGTATGCCGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGGTATGCCGTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGTCAGACTGTCGATGAAGCCCTGAAAGAAGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGACTGTCGATGAAGCCCTGAAAGAAGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
```

Search completed: August 1, 2005, 04:28:44
Job time : 1101.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacacacgaatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3300	100.0	3300	6	AX202428	Sequence
2	3216.8	97.5	3300	6	AX202430	Sequence
3	3207.2	97.2	3300	6	AX202426	Sequence
4	3207.2	97.2	3300	6	AX202427	Sequence
5	3207.2	97.2	3300	6	AX202429	Sequence
6	2705.8	82.0	7173	6	AX284144	Sequence
7	2705.8	82.0	8101	6	AX172306	Sequence
8	2701	81.8	6648	6	AX377531	Sequence
9	2700	81.8	7475	6	A91965	Sequence 1
10	2700	81.8	7475	6	AR031992	Sequence
11	2700	81.8	7475	6	AR207294	Sequence
12	2698	81.8	9191	6	AX377532	Sequence
13	2676.2	81.1	6724	12	AF097412	Sequence
14	2636.8	79.9	6806	6	AX378208	Sequence
15	2631	79.7	6706	12	AF031088	Sequence
16	2598.2	78.7	6748	12	AF031813	Sequence
17	1609.4	48.8	5558	6	CQ794769	Sequence
18	1474	44.7	4700	6	CQ846980	Sequence
19	1462.4	44.3	4921	6	AX768201	Sequence

20	1462.4	44.3	4935	6	AX768200	Sequence
21	1462.4	44.3	4945	6	AX768202	Sequence
22	1462.4	44.3	4951	6	AX768203	Sequence
c 23	1432.4	43.4	6755	12	AF140577	Integrati
24	1417.6	43.0	1511	6	AX431576	Sequence
25	1417.6	43.0	4297	12	AV219685	Expressio
26	1417.6	43.0	4359	12	AV219686	Expressio
27	1417.6	43.0	4419	12	AV219688	Expressio
28	1417.6	43.0	4855	12	AV219689	Expressio
29	1417.6	43.0	4968	12	AV219687	Expressio
30	1417.6	43.0	6956	12	AV219682	Shuttle e
31	1417.6	43.0	7018	12	AV219683	Shuttle e
32	1417.6	43.0	7509	12	AV219684	Shuttle e
33	1400.8	42.4	4566	12	AF177933	Cloning v
34	1400.8	42.4	4635	12	AF177932	Cloning v
35	1400.8	42.4	6399	12	AY302759	Expressio
36	1400.8	42.4	6587	12	AY289542	Expressio
37	1362	41.3	5201	6	AR493834	Sequence
38	1362	41.3	5201	6	AR493835	Sequence
39	1329.2	40.3	4806	12	AF050464	Expressio
c 40	1238.4	37.5	4803	12	AY243506	Cloning v
c 41	1238.4	37.5	5024	6	CQ846977	Sequence
42	1237.2	37.5	1922	6	AX478528	Sequence
43	1204	36.5	4506	12	AB116367	Expressio
44	1204	36.5	4557	6	EL3740	DNA sequen
45	1204	36.5	4557	6	AR373251	Sequence

ALIGNMENTS

RESULT 1
AX202428
LOCUS AX202428 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 16 from Patent WO0152620.
ACCESSION AX202428
VERSION AX202428.1 GI:15392176
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 16 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m3 and zinc finger protein ZFPm3"

ORIGIN	Query Match	100.0%;	Score 3300;	DB 6;	Length 3300;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3300;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	CGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGCGCATGATGAGCCCGGAAGAGA	60	
Db	1	CGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGCGCATGATGAGCCCGGAAGAGA	60	
Qy	61	GTCATTTCAGGTGGTGAATGTGA	AAACCGTAACTTATACGATGTCGACAGTATGCCG	120	
Db	61	GTCATTTCAGGTGGTGAATGTGA	AAACCGTAACTTATACGATGTCGACAGTATGCCG	120	
Qy	121	GTGTCCTTATCAGACCGTTTCC	CGCGGTGGTGAACCCAGGCCACCGCTTTCTGCGAAA	180	
Db	121	GTGTCCTTATCAGACCGTTTCC	CGCGGTGGTGAACCCAGGCCACCGCTTTCTGCGAAA	180	
Qy	181	CGCGGAAAAAGTGGAACGGCG	GATGGCGAGCTGAATTACATCCCAACCGGTGGCAC	240	

181	CGCGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAAATTACATTCCTCAACCGCTGGCAC	240
QY	AACAACTGCGCGGCAAAACAGTCTGTTCTGATTGGCGTGTGCCACCTCCAGTCTGGCCCTGC	300
Db	AACAACTGCGCGGCAAAACAGTCTGTTGCTGATTGGCGTGGCACTTCCAGTCTGGCCCTGC	300
QY	ACGGCCGTCGCAAAATTGTTCGGCGGCGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG	360
Db	ACGGCCGTCGCAAAATTGTTCGGCGGCGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG	360
QY	TGSGTGGTTCGATGGTGAACGAAAGCGCGTCGAAAGCTGTAAAGCGGCGGTGCACAATC	420
Db	TGSGTGGTTCGATGGTGAACGAAAGCGCGTCGAAAGCTGTAAAGCGGCGGTGCACAATC	420
QY	TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAACATCCGCTGGATCACCAAGATGCCA	480
Db	TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAACATCCGCTGGATCACCAAGATGCCA	480
QY	TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTCTTCGATGCTCTGACCAAG	540
Db	TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTCTTCGATGCTCTGACCAAG	540
QY	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGATCGGCGGTGCAGCATC	600
Db	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGATCGGCGGTGCAGCATC	600
QY	TGSGTCGATTGGGTCACAGCAAAATCGCGCTGTAGCGGGGCCATTAAAGTTCGTCTCGG	660
Db	TGSGTCGATTGGGTCACAGCAAAATCGCGCTGTAGCGGGGCCATTAAAGTTCGTCTCGG	660
QY	CGCGTCTCGGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	CGCGTCTCGGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	CGGAAACGGGAAGCGCATGAGTGCCTATCTCCGTTTTTCAACAAACCATGCAAAATGCTGA	780
Db	CGGAAACGGGAAGCGCATGAGTGCCTATCTCCGTTTTTCAACAAACCATGCAAAATGCTGA	780
QY	ATGAGGCAATCGTTTCCCACTCGCATGTGTTGTGCCAAGATCAGATGCGCGTGGCGCGAA	840
Db	ATGAGGCAATCGTTTCCCACTCGCATGTGTTGTGCCAAGATCAGATGCGCGTGGCGCGAA	840
QY	TGCGCGCATTACCGAGTCCGGGCTCGCGTGTGTGTCGGAATCTCGGTAGTGGGATACG	900
Db	TGCGCGCATTACCGAGTCCGGGCTCGCGTGTGTGTCGGAATCTCGGTAGTGGGATACG	900
QY	ACGATACCGAAGACAGCTCATGTTTATATCCGCGGTTAAACACCATCAAAACAGGATTTTC	960
Db	ACGATACCGAAGACAGCTCATGTTTATATCCGCGGTTAAACACCATCAAAACAGGATTTTC	960
QY	GCCTGTGCGGGCAAAACAGCGTGGACCGCTGTGTCGAATCTCTCAGGCGCAGCGGTGA	1020
Db	GCCTGTGCGGGCAAAACAGCGTGGACCGCTGTGTCGAATCTCTCAGGCGCAGCGGTGA	1020
QY	AGGGCAATCAGCTGTTCGCCGTCTCACTGTGTGAAGAAAGAAACACCCCTCGCGCCCAATA	1080
Db	AGGGCAATCAGCTGTTCGCCGTCTCACTGTGTGAAGAAAGAAACACCCCTCGCGCCCAATA	1080
QY	CGAAACCGGCTCTCCCGCGGTTGCGCGATTCATTAATGCAGCTGGCACGACAGGTTT	1140
Db	CGAAACCGGCTCTCCCGCGGTTGCGCGATTCATTAATGCAGCTGGCACGACAGGTTT	1140
QY	CCGCACTGGAAAGCGGCACTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Db	CCGCACTGGAAAGCGGCACTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
QY	GCACAAATTCATGTTTGACAGCTTATCATCTGCACTGCAACGCGTGCACAAATGCTTCTGGCG	1260
Db	GCACAAATTCATGTTTGACAGCTTATCATCTGCACTGCAACGCGTGCACAAATGCTTCTGGCG	1260
QY	TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTAAATCACTGCATAAATTCG	1320

Db	1261	TCAGCAGCCATCGAAGAGCTGTGTATGGCTGTGAGGTGCTGTAATCACTGCATAATTCG	1330
Qy	1321	TGTCGCTCAAGCGCACATCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCATAAACGGTT	1380
Db	1321	TGTGCTTCAAGCGCACATCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCATAAACGGTT	1380
Qy	1381	CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Db	1381	CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Qy	1441	ATTGTAGCGGATACCAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTTACGA	1500
Db	1441	ATTGTAGCGGATACCAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTTACGA	1500
Qy	1501	GCATTCCACCAAGGACCATAGATTATGAAAACTGAAGAAGTAAACTGGGTAATCTGG	1560
Db	1501	GCATTCCACCAAGGACCATAGATTATGAAAACTGAAGAAGTAAACTGGTAACTCG	1560
Qy	1561	ATTAAACGGCGATAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAGAT	1620
Db	1561	ATTAAACGGCGATAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAGAT	1620
Qy	1621	ACCGGAATTTAAAGTACACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGTT	1680
Db	1621	ACCGGAATTTAAAGTACACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGTT	1680
Qy	1681	CGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACGCTTGGTGCTAC	1740
Db	1681	CGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACGCTTGGTGCTAC	1740
Qy	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGAGACAGCTGAT	1800
Db	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGAGACAGCTGAT	1800
Qy	1801	CCGTTTACTGGGATGCGTACGTTTAAACCGGCAAGCTGATTCCTTACCAGTCGCTGTT	1860
Db	1801	CCGTTTACTGGGATGCGTACGTTTAAACCGGCAAGCTGATTCCTTACCAGTCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGAA	1920
Qy	1921	GAGATCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC	1980
Qy	1981	CTGCAAGAACGGTACTTCACCTGGCGCTGATTGCTGCTGAACGGGGTTATCGGTTCAAG	2040
Db	1981	CTGCAAGAACGGTACTTCACCTGGCGCTGATTGCTGCTGAACGGGGTTATCGGTTCAAG	2040
Qy	2041	TATGAAACCGGCAAGTACGACATTTAAAGACGTGGGCGTGGATAAACGCTCGCGGAAAGCG	2100
Db	2041	TATGAAACCGGCAAGTACGACATTTAAAGACGTGGGCGTGGATAAACGCTCGCGGAAAGCG	2100
Qy	2101	GGTCTGACCTTCCTGGTTGACCTGATTAAACCAACACATGAATGCAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCCTGGTTGACCTGATTAAACCAACACATGAATGCAGACACCGATTAC	2160
Qy	2161	TCCATCGCAAGCTGCTTTTATAAAGCGAAACAGCGATGACCATCAACGCGCCGTGG	2220
Db	2161	TCCATCGCAAGCTGCTTTTATAAAGCGGAAACAGCGATGACCATCAACGCGCCGTGG	2220
Qy	2221	GCATGFTCCAAATCGACACCCAGCAAAAGTGAATTAATGGTGTAAACCGGTACTGCCGACCTTC	2280
Db	2221	GCATGFTCCAAATCGACACCCAGCAAAAGTGAATTAATGGTGTAAACCGGTACTGCCGACCTTC	2280
Qy	2281	AAGGTCACCATTCGAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCCGAGT	2340
Db	2281	AAGGTCACCATTCGAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCCGAGT	2340
Qy	2341	CCGAAACAAAGAGCTGGCAAAAGTTCTTCGAAAACTATCTCCTGACATGATGAAGTCTG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAAGTTCTTCGAAAACTATCTCCTGACATGATGAAGTCTG	2400

Qy	841	TGCGCGCCATTACCGAGTC	CGGGCTCGCGT	TGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCCATTACCGAGTC	CGGGCTCGCGT	TGGTGGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGGTTAAACACCATCAAAACAGGATTTTC	960		
Db	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGGTTAAACACCATCAAAACAGGATTTTC	960		
Qy	961	GCCTGCTGGGGCAAAACACAGCGTGAACCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020		
Db	961	GCCTGCTGGGGCAAAACACAGCGTGAACCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020		
Qy	1021	AGGCAATCAGCTGTGTCGCCGTCTCACTTGGTGAAGAAACACACCTGCGCCCAATA	1080		
Db	1021	AGGCAATCAGCTGTGTCGCCGTCTCACTTGGTGAAGAAACACACCTGCGCCCAATA	1080		
Qy	1081	CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGCGTTT	1140		
Db	1081	CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGCGTTT	1140		
Qy	1141	CCGACCTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCATCATAG	1200		
Db	1141	CCGACCTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCATCATAG	1200		
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACGTGACCGTGACCAATGCTTCTCGCG	1260		
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACGTGACCGTGACCAATGCTTCTCGCG	1260		
Qy	1261	TCAGGCAGCCATCGGAAGCTGTGATGTGGCTGTGCAGTGGTAAATCACTGCATAATTCG	1320		
Db	1261	TCAGGCAGCCATCGGAAGCTGTGATGTGGCTGTGCAGTGGTAAATCACTGCATAATTCG	1320		
Qy	1321	TGTCGCTCAAGGCGCACTCCGGTCTCGATAATGTTTTTGGCCGACATCAATAACGGTT	1380		
Db	1321	TGTCGCTCAAGGCGCACTCCGGTCTCGATAATGTTTTTGGCCGACATCAATAACGGTT	1380		
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440		
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440		
Qy	1441	ATTGTAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTTCACGA	1500		
Db	1441	ATTGTAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTTCACGA	1500		
Qy	1501	GCACTTACCAAACAGGACCATAGATTGAAAATCTGAAAGAGTAACTGGTAAATCTCG	1560		
Db	1501	GCACTTACCAAACAGGACCATAGATTGAAAATCTGAAAGAGTAACTGGTAAATCTCG	1560		
Qy	1561	ATTAAACGGCGATAAAGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAGAT	1620		
Db	1561	ATTAAACGGCGATAAAGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAGAT	1620		
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATCCCAACAGTT	1680		
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATCCCAACAGTT	1680		
Qy	1681	GCGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCAACGACGCTTGGTGGCTAC	1740		
Db	1681	GCGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCAACGACGCTTGGTGGCTAC	1740		
Qy	1741	GCTCAATCTGGCCGTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT	1800		
Db	1741	GCTCAATCTGGCCGTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT	1800		
Qy	1801	CGGTTTACTGGGATCCCGTACGTTTAAACGGCAAGCTGATTTGCTTACCCGATCGCTGTT	1860		
Db	1801	CGGTTTACTGGGATCCCGTACGTTTAAACGGCAAGCTGATTTGCTTACCCGATCGCTGTT	1860		
Qy	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCGCGAACCCCGCAAAAACCTCGGAA	1920		
Db	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCGCGAACCCCGCAAAAACCTCGGAA	1920		

QY	1921	GAGATCCGGCGCGTGGATAAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGATGTTTCAAC	1981
DB	1921	GAGATCCGGCGCGTGGATAAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGATGTTTCAAC	1980
QY	1981	CTGCAAGAAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAACGGCAAGTAGACATATAAGACGCTGGGCGTGGATAACGCTGCGCGAAAGCG	2100
DB	2041	TATGAAAACGGCAAGTAGACATATAAGACGCTGGGCGTGGATAACGCTGCGCGAAAGCG	2100
QY	2101	GGTCTGACCTTTCCTGGTTGACCTGATTTAAAAAABAACATGATGATGACACACCGGATTAC	2160
DB	2101	GGTCTGACCTTTCCTGGTTGACCTGATTTAAAAAABAACATGATGATGACACACCGGATTAC	2160
QY	2161	TCCATCGCAAGAGTGCCTTTTAATAAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG	2220
DB	2161	TCCATCGCAAGAGTGCCTTTTAATAAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG	2220
QY	2221	GCATGTTCCAACTCGACACACGCAAGTGAATTATGTTGTTAAACGCTACTGCGGACCTTC	2280
DB	2221	GCATGTTCCAACTCGACACACGCAAGTGAATTATGTTGTTAAACGCTACTGCGGACCTTC	2280
QY	2281	AAGGGTCAACCATCTCCAAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTAAACGCGCGACT	2340
DB	2281	AAGGGTCAACCATCTCCAAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTAAACGCGCGACT	2340
QY	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTTCCTCGAAAACTATCTGCTGATGATGAAAGTCTG	2400
DB	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTTCCTCGAAAACTATCTGCTGATGATGAAAGTCTG	2400
QY	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGACGCTTACGAGGAGAG	2460
DB	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGACGCTTACGAGGAGAG	2460
QY	2461	TTGGCGAAAGATCCAGTATTCGCCGACCATGGAACCGCCAGAAAGGTGAAATCATG	2520
DB	2461	TTGGCGAAAGATCCAGTATTCGCCGACCATGGAACCGCCAGAAAGGTGAAATCATG	2520
QY	2521	CCGAAACATCCCGCAGATGTCGCGTTTCGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
DB	2521	CCGAAACATCCCGCAGATGTCGCGTTTCGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
QY	2581	GCCAGCGGTCGTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
DB	2581	GCCAGCGGTCGTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
QY	2641	AACAAACAACAATTAACAATAACAACACTCGGGATCGAGGGAAGGATTCAGAAATTC	2700
DB	2641	AACAAACAACAATTAACAATAACAACACTCGGGATCGAGGGAAGGATTCAGAAATTC	2700
QY	2701	GGATCTCTTCTCTGTGGCCAGCGCGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT	2760
DB	2701	GGATCTCTTCTCTGTGGCCAGCGCGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT	2760
QY	2761	CCGGAATGTGGTAAGTCTTTCAGCGATCTGCGCCACTGGTTCGCCACGAGCGTACCCAC	2820
DB	2761	CCGGAATGTGGTAAGTCTTTCAGCGAGCAGTCCCTGCTGGTCGCCACGAGCGTACCCAC	2820
QY	2821	ACGGGTGAAAAACCGTATAAATGCCAGATGCGGCAATCTTTTAGCACGAGCGGCTCC	2880
DB	2821	ACGGGTGAAAAACCGTATAAATGCCAGATGCGGCAATCTTTTAGCACGAGCGGCTCC	2880
QY	2881	CTGGTCGCCCATCAACGCACTCATACTGCGGAGAGCCATACAAATGTCAGAAATGTGGC	2940
DB	2881	CTGGTCGCCCATCAACGCACTCATACTGCGGAGAGCCATACAAATGTCAGAAATGTGGC	2940
QY	2941	AAGTCTTTTACGCGAGCTTCGAGCTGTCGCGCACCAACGCTACTCTCAACCGGGAGAG	3000
DB	2941	AAGTCTTTTACGCGAGCTTCGAGCAACCTGGTGGGCCACCAACGCTACTCTCAACCGGGAGAG	3000
QY	3001	CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTTCAGCCAGAGCAGCTCCCTGCTGCGCCAC	3060

Db 3001 |||||||CCCTATGCTTGTCCGGAATGTTGAAGTCCCTTTCAGCACACAGTGGCTCTCTGGTTAGACAC 3060
Qy 3061 CAGCGTACCCACACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGT 3120
Db 3061 CAGCGTACCCACACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Qy 3121 GACTGCCGGACCTTGTCTGGCCATCAACGCACTCATCTATCTGGCGAGAACCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTTGTGAAACGGCCATCAACGCACTCATCTATCTGGCGAGAACCATACAAATGT 3180
Qy 3181 CCAGAATGTGCAAGTCTTCTCCCAATCCAGCCATCTCGTCCGCGACCAAGTACTCAC 3240
Db 3181 CCAGAATGTGCAAGTCTTCTCAACTTCAGCAACTTGGTCCGTCACCAAGTACTCAC 3240
Qy 3241 ACCGTAATAAAAACTAGTGGCCAGCGCCGACAGTACCGGTACAGCTTCCGAGCTACGCT 3300
Db 3241 ACCGTAATAAAAACTAGTGGCCAGCGCCGACAGTACCGGTACAGCTTCCGAGCTACGCT 3300

RESULT 3
AX202426
LOCUS AX202426 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 14 from Patent WO0152620.
ACCESSION AX202426
VERSION AX202426.1 GI:15392173
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 14 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m1 and zinc finger protein
ZFpml"

ORIGIN
Query Match 97.2%; Score 3207.2; DB 6; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGCGATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGCGATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCATTCAGGGTGTGAATGTGAACACCAAGTATACGATTCGAGAGTATGCG 120
Db 61 GTCATTCAGGGTGTGAATGTGAACACCAAGTATACGATTCGAGAGTATGCG 120
Qy 121 GTGTCCTTTATCAGACCGTTTCCCGGTGGTGAACAGCGACGACGTTTCTGCGAAAA 180
Db 121 GTGTCCTTTATCAGACCGTTTCCCGGTGGTGAACAGCGACGACGTTTCTGCGAAAA 180
Qy 181 CCGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCCAAACCGCTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCCAAACCGCTGGCAC 240
Qy 241 AACAACTGGCGGGCAAAACAGTCTGTCTGATTTGGGTGTCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTCTGATTTGGGTGTCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCCCGTCCGCAATTTGTCGGCGGATTAATCTCGCGCCGATCACTGGGTGCCAGCG 360
Db 301 ACGCCCGTCCGCAATTTGTCGGCGGATTAATCTCGCGCCGATCACTGGGTGCCAGCG 360

Qy 361 TGGTGGTGTTCATGGTAGAACGAAGCGCGCTGGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTTCATGGTAGAACGAAGCGCGCTGGAAGCCTGTAAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGGCAACCGCTCAGTGGGCTGATCATTAATACTATCCGTGATCACCAAGATGCCA 480
Db 421 TTCTCGGCAACCGCTCAGTGGGCTGATCATTAATACTATCCGTGATCACCAAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGCAGATTCGGGTACCCAGCAAAATCGGCGTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db 601 TGGTGCAGATTCGGGTACCCAGCAAAATCGGCGTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCGGGTTTCAACAAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCGGGTTTCAACAAACCAATGCAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGCCAAGCATCAGATGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGCCAAGCATCAGATGCGCTGGCGCGAA 840
Qy 841 TGGCGCCCATTAACCGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGTAGTGGGATAG 900
Db 841 TGGCGCCCATTAACCGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACCAATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACCAATCAACAGGATTTTC 960
Qy 961 GCCTGTGGGGCAAAACAGCGGTGGACCGCTTGTGTCGAATCTCTCAGGCGCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACAGCGGTGGACCGCTTGTGTCGAATCTCTCAGGCGCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGTGTGAAAAAGAAAAACCACTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGTGTGAAAAAGAAAAACCACTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCAGAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCAGAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGCAAGTGTGCAAGTGTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGCAAGTGTGCAAGTGTCTGGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGGTGTGAGTGTGAGTGTGAGTGTG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGGTGTGAGTGTGAGTGTGAGTGTG 1320
Qy 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500

ORIGIN	ZFPm2"	Db	961
Query Match	97.2%; Score 3207.2; DB 6; Length 3300;		
Best Local Similarity	98.2%; Pred. No. 0;	QY	1021
Matches 3242; Conservative	0; Mismatches 58; Indels 0; Gaps 0;	Db	1021
QY	1	QY	1081
Db	1	Db	1081
QY	61	QY	1141
Db	61	Db	1141
QY	121	QY	1201
Db	121	Db	1201
QY	181	QY	1321
Db	181	Db	1321
QY	241	QY	1381
Db	241	Db	1381
QY	301	QY	1441
Db	301	Db	1441
QY	361	QY	1501
Db	361	Db	1501
QY	421	QY	1561
Db	421	Db	1561
QY	481	QY	1621
Db	481	Db	1621
QY	541	QY	1681
Db	541	Db	1681
QY	601	QY	1741
Db	601	Db	1741
QY	661	QY	1801
Db	661	Db	1801
QY	721	QY	1861
Db	721	Db	1861
QY	781	QY	1921
Db	781	Db	1921
QY	841	QY	1981
Db	841	Db	1981
QY	901	QY	2041
Db	901	Db	2041
QY	961	QY	2101

```
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACACATGAATGACAGACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACACATGAATGACAGACCGATTAC 2160
Qy 2161 TCCATCGCAGAGTGCCTTTAATAAGCGCAACAGAGATGACCATCAACCGGCCGTGG 2220
Db 2161 TCCATCGCAGAGTGCCTTTAATAAGCGCAACAGAGATGACCATCAACCGGCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACAGCAAGAGTGAATATGTTGTAACGGTACTCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGAGTGAATATGTTGTAACGGTACTCCGACCTTC 2280
Qy 2281 AAGGTCACACCATCAACACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACCGCCGACGT 2340
Db 2281 AAGGTCACACCATCAACACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACCGCCGACGT 2340
Qy 2341 CCGAACAAAGAGTGGGCAAAAGAGTTCCTCGAAACATCTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CCGAACAAAGAGTGGGCAAAAGAGTTCCTCGAAACATCTATCTGCTGACTGATGAAGGCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Qy 2461 TTGGCGAAAGATCCACGATATTCGCGCCACCATGGAAACCGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATATTCGCGCCACCATGGAAACCGCCGAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGCGTTTCTGTTGATGCGGTGCGTATCTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCGTTTCTGTTGATGCGGTGCGTATCTGCGGTGATCAAGCC 2580
Qy 2581 GCCAGCGTCTGACAGTCTCGATGAGCGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTCGATGAGCGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCTCTTCTCTGTGGCCAGCGGCGCTTCGAGCGCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGGCGCTTCGAGCGCGGGGAGAGCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTCTCAGCGATCTCTGCGCACCTGTTGCGCCACCGAGCTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTCTCAGAGCTCTCAGCTGTCTCAGCTGTGCGCCACCGAGCTACCCAC 2820
Qy 2821 ACGGTTGAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCACCGCGGCTCC 2880
Db 2821 ACGGTTGAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGCATCTACTTGGCGAGAGGCGCATACAATGTCCGAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCATCTACTTGGCGAGAGGCGCATACAATGTCTCAGAAATGTGGC 2940
Qy 2941 AAGTCTTTTACGCCAGAGCTCCAGCGCTGTGGCGCCACCAAGCTACTCACCGGGGAGAAG 3000
Db 2941 AAGTCTTTTCTCGTCTGACAAATCTGTTCGGGACCAACGCTACTCACCGGGGAGAAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCCAGAGCAGCTCCCTGGTGGCGCAC 3060
Db 3001 CCCTATGCTTGTTCGGAATGTGGTAAGTCTTTCAGCCGCGCAGATACCTGGTGGCGCAC 3060
Qy 3061 CAGGTTACCCACAGCGGTGAACCCGATATAAATGCCAGAGTGGCGCAATCTTTTAGT 3120
Db 3061 CAGGTTACCCACAGCGGTGAACCCGATATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
Qy 3121 GACTGCGCGCACCTTGTTCGCCATCAACGCACTCATATCTGCGGAGAGGCCATACAATGT 3180
Db 3121 CAGCGCGCACCTTGTTCGCCATCAACGCACTCATATCTGCGGAGAGGCCATACAATGT 3180
```

```
Qy 3181 CCAGAATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTGTCGGGACCAACAGTACTCAC 3240
Db 3181 CCAGAATGTGGCAAGTCTTTCTCCCAATCTGCTCGGTCTGACAAATCTGTCGGGACCAACAGTACTCAC 3240
Qy 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCCTGACGACGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCCTGACGACGTTCGGGACTACGCT 3300

RESULT 5
AX202429 LOCUS AX202429 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 17 from Patent WO0152620.
ACCESSION AX202429
VERSION AX202429.1 GI:15392177
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 17 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source 1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Partial sequence of pMal-m4 and zinc finger protein ZFPm4"

ORIGIN
Query Match 97.2%; Score 3207.2; DB 6; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 CCACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 1 CCACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTCAGGTTGGTGAATGTGAACCACTGATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGTTGGTGAATGTGAACCACTGATACGATGTGCGAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGAGCCAGCCAGCTTTCTGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGAGCCAGCCAGCTTTCTGCGAAA 180
Qy 181 CCGGGGAAAAGTGGAGCGCGATGCGGAGCTGAATTAACATTCCCAACCGGTGGCAC 240
Db 181 CCGGGGAAAAGTGGAGCGCGATGCGGAGCTGAATTAACATTCCCAACCGGTGGCAC 240
Qy 241 AACCACTGGCGGGCAACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Db 241 AACCACTGGCGGGCAACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTCCGAAATTTGTGCGCGGATTAATATCTCGCGCCGATCACTGGGTGCCAGCG 360
Db 301 ACGCGCGTCCGAAATTTGTGCGCGGATTAATATCTCGCGCCGATCACTGGGTGCCAGCG 360
Qy 361 TGTGTGTGTCGATGGTAGAACGAGCGGTGCAAGCGCTGTAAGCGGGTGCACAATC 420
Db 361 TGTGTGTGTCGATGGTAGAACGAGCGGTGCAAGCGCTGTAAGCGGGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATATCCGCTGGATGACAGAGTGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATATCCGCTGGATGACAGAGTGCCA 480
Qy 481 TTGCTGTGGAGAGTGGCTGCACTAAATGTTCCGGCGTATTTCTTGATGTCTTGACGAGA 540
Db 481 TTGCTGTGGAGAGTGGCTGCACTAAATGTTCCGGCGTATTTCTTGATGTCTTGACGAGA 540
```

QY	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
QY	601	TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTTAAGTCTGTCTCGG	660
DB	601	TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTTAAGTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCACGCGATAG	720
DB	661	CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCACGCGATAG	720
QY	721	CGGAACGGGAAGCGACCTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGCGACCTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTCCCACTCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGGCA	840
DB	781	ATGAGGGCATCGTCCCACTCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGGCA	840
QY	841	TGCGCGCCATACCGAGTCCGGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGGATAG	900
DB	841	TGCGCGCCATACCGAGTCCGGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGGATAG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAAAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAAAACAGGATTTTC	960
QY	961	GCCTCTGGGGCAACACAGCTGGACCGCTCTGCACTCTCAGGCGCCAGGCGGTGA	1020
DB	961	GCCTCTGGGGCAACACAGCTGGACCGCTCTGCACTCTCAGGCGCCAGGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA	1080
DB	1021	AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGACGTGGCACAGGATTTTC	1140
DB	1081	CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGACGTGGCACAGGATTTTC	1140
QY	1141	CCCGACTGGAAAGCGGAGTGGCGCAACGCAATTAATGAGTTAGCTCACTCATTAG	1200
DB	1141	CCCGACTGGAAAGCGGAGTGGCGCAACGCAATTAATGAGTTAGCTCACTCATTAG	1200
QY	1201	GCACAAATCTCATGTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATTAATTCG	1320
DB	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATTAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCCGACATCAACGGTT	1380
DB	1321	TGTCGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCCGACATCAACGGTT	1380
QY	1381	CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCACTCGGCTCGTAAATGTTGGA	1440
DB	1381	CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCACTCGGCTCGTAAATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA	1500
DB	1441	ATTGTGAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA	1500
QY	1501	GCACCTCACCAACAGGACATAGATTATGAACCTGGAAGAGGTAACTGTAATCTCG	1560
DB	1501	GCACCTCACCAACAGGACATAGATTATGAACCTGGAAGAGGTAACTGTAATCTCG	1560
QY	1561	ATTAAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAGAT	1620
DB	1561	ATTAAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAGAT	1620

QY	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGNTAACTGGAGAGAAATTTCCACAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGNTAACTGGAGAGAAATTTCCACAGGTT	1680
QY	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
DB	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT	1800
DB	1741	GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT	1860
DB	1801	CGGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATTTCAAC	1980
DB	1921	GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATAACGCTGGCGGAAAGCG	2100
DB	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATAACGCTGGCGGAAAGCG	2100
QY	2101	GCTCTGACCTTCTGCTGTTGACCTGATTAATAAAGCGGAAACAGCGATGACCATCAACCGGCTGG	2160
DB	2101	GCTCTGACCTTCTGCTGTTGACCTGATTAATAAAGCGGAAACAGCGATGACCATCAACCGGCTGG	2160
QY	2161	TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACCGGCTGG	2220
DB	2161	TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACCGGCTGG	2220
QY	2221	GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGTTAAACGCTACTGCGGACCTTC	2280
DB	2221	GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGTTAAACGCTACTGCGGACCTTC	2280
QY	2281	AAGGTCACCAATCAAAACCGTTCTGCGCTGCTGAGCGCAGGTATTAACGCCGCCAGT	2340
DB	2281	AAGGTCACCAATCAAAACCGTTCTGCGCTGCTGAGCGCAGGTATTAACGCCGCCAGT	2340
QY	2341	CGGAACAAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG	2400
DB	2341	CGGAACAAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG	2400
QY	2401	GAAGCGGTTAATAAGACAAACCGCTGGCTGGCTAGCGCTGAACTCTTACGAGGAAGAG	2460
DB	2401	GAAGCGGTTAATAAGACAAACCGCTGGCTGGCTAGCGCTGAACTCTTACGAGGAAGAG	2460
QY	2461	TTGGCGAAAGATCCAGTATTGGCCGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG	2520
DB	2461	TTGGCGAAAGATCCAGTATTGGCCGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG	2520
QY	2521	CGGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC	2580
DB	2521	CGGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC	2580
QY	2581	GCACGCGGTCTGACAGCTGTCGATGAAGCTGGAAGACGCGCAGACTAATTCGAGCTCG	2640
DB	2581	GCACGCGGTCTGACAGCTGTCGATGAAGCTGGAAGACGCGCAGACTAATTCGAGCTCG	2640
QY	2641	AACAAACAAACAAATAACAACTTCGGGATCGAGGGAGGATTTTCAAGATTC	2700
DB	2641	AACAAACAAACAAATAACAACTTCGGGATCGAGGGAGGATTTTCAAGATTC	2700
QY	2701	GGATCCTCTTCTCTGCTGGCCACGCGGCTCTGAGAGCCCGGGGAGAACCCCTATGCTGT	2760

Db 2701 GGATCTCTTCTCTGTGCGCCAGCGGCCCTCGAGCCCGGAGAGCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTCTTACGCGATCTGCGCACCTGGTTCGCGCACCGAGGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTCTTACGCGAGCAGCTGCTGCGCGCACCGAGGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCACACGCGCTCC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCACAGCAGCAGC 2880
Qy 2881 CTGGTGGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Qy 2941 AAGTCTTTTCAGCGAGGTCCAGCGCTGGTGGCGCACCAACGTAAGTCTTACCGAGAGAG 3000
Db 2941 AAGTCTTTTCAGGATGTTGCGTATCTTGGGAGGACCAAGTACTCACCGGGGAGAG 3000
Qy 3001 CCCTATGCTTGTCCCGAATGTGGTAAGTCTTTCAGCGAGAGCAGCTCCCTGTCGCGCAC 3060
Db 3001 CCCTATGCTTGTCCCGAATGTGGTAAGTCTTTCAGCGAGAGCAGCTCCCTGTCGCGCAC 3060
Qy 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGT 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGT 3120
Qy 3121 GACTGCGCGACCTTGTCTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 GCGAGCGATAACCTGTTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Qy 3181 CCAGATGTGCAAGTCTTCTCCCAATCGAGCCATCTGTCGCGCACCAAGTACTCAC 3240
Db 3181 CCAGATGTGCAAGTCTTCTCCCAATCGAGCCATCTGTCGCGCACCAAGTACTCAC 3240
Qy 3241 ACCGTTAAAAAACTAGTGGCGAGCGCCGAGTACCGTACGAGTTCGCGAGTACGCT 3300
Db 3241 ACCGTTAAAAAACTAGTGGCGAGCGCCGAGTACCGTACGAGTTCGCGAGTACGCT 3300

RESULT 6
AX284144
LOCUS AX284144 7373 bp DNA linear PAT 20-NOV-2001
DEFINITION Sequence 1 from Patent WO0178756.
ACCESSION AX284144
VERSION AX284144.1 GI:17044832

KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
Wiederanders, B. and Maubach, G.
Agent for postoperative use after the removal of bone tumours
Patent: WO 0178756-A 1 25-OCT-2001;
Depuy Biotech Jena GmbH (DE)

FEATURES
source
1. 7373
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Expressionsvektor pMalc2cchbnp2"
1528. 3453
/note="unnamed protein product; /produkt=Fusionsprotein
MBP Cystatin C BMP-2"
/codon_start=1
/transl_table=11
/protein_id="CAD12160.1"
/db_xref="GI:17044833"
/translation="MKTEEGKLVIWINGDKYNGLAEVGKKFEKDTGKIVTVEHPDKL
EEKFPQVAATGDDPDIIPWAHDPFGVAQSGLLAEITPDKAFQDKLYPFFWDAVRNG
KLAYPAVATGALSILYINKDLINPPKPIWEEIPALDKELKAGKSALENLQEPYFTWP
LIAADGGYAFKYENGKDIKDVGVNDAGAGLFLVDLKNHMDYDIAEAAAF
NKGETAMTINGPWAWSNIDTSKVNIGYVTLPTFKGQPSKPFVGLSAGINAAASPNKEL

CDS

misc_feature 2689
/note="Faktor Xa Protease /position=388 (AMINO
ACIDsequenz)"
3067. 3108
misc_feature
/note="/spacer zwischen Cystatin C und BMP-2"
ORIGIN
Query Match 82.08; Score 2705.8; DB 6; Length 7373;
Best Local Similarity 99.24; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATATAGCGCCCGGAGAGA 60.
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATATAGCGCCCGGAGAGA 60
Qy 61 GTCAATTCAGGTTGGTGAATGTGAACACGTAACGTTATACGATGTCCGAGAGTATCCG 120
Db 61 GTCAATTCAGGTTGGTGAATGTGAACACGTAACGTTATACGATGTCCGAGAGTATCCG 120
Qy 121 GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCCAGGTTTTCGCGAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCCAGGTTTTCGCGAAA 180
Qy 181 CCGCGGAAAAAGTGGAGCGCGCATGGCGAGCTGAATTTACATTCCCAACCGCGTGGCAC 240
Db 181 CCGCGGAAAAAGTGGAGCGCGCATGGCGAGCTGAATTTACATTCCCAACCGCGTGGCAC 240
Qy 241 AACCACTGGCGGCGCAACAGTCTGTTGCTGATTTGGCGTTGCGCACTCCTCAGTCTGGCCTGC 300
Db 241 AACCACTGGCGGCGCAACAGTCTGTTGCTGATTTGGCGTTGCGCACTCCTCAGTCTGGCCTGC 300
Qy 301 ACCGCGCGTCCGCAAAATTTGTCGCGCGCATTAATAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 301 ACCGCGCGTCCGCAAAATTTGTCGCGCGCATTAATAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Qy 361 TGTGTTGTCGATGGTAGAAACGAGCGCGTGAAGCCTGTAAAGCGGCGGTGCACAAATC 420
Db 361 TGTGTTGTCGATGGTAGAAACGAGCGCGTGAAGCCTGTAAAGCGGCGGTGCACAAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
Qy 481 TTCTGTGGAAGCTGCTGCATTAATTTTCGCGCGTTATTTCTTGATGTCTCTGACAGA 540
Db 481 TTCTGTGGAAGCTGCTGCATTAATTTTCGCGCGTTATTTCTTGATGTCTCTGACAGA 540
Qy 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGGACCTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGGACCTGGGCGTGGAGCATC 600
Qy 601 TGTGTCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGCTCCG 660
Db 601 TGTGTCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGCTCCG 660
Qy 661 CCGCTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGCTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CCGAACCGGAAGCGGCTGGAGTGGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
Db 721 CCGAACCGGAAGCGGCTGGAGTGGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGCGCATCGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGCGCATCGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840


```
QY 841 TCGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATTACCGAAGACAGCTCATGTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATTACCGAAGACAGCTCATGTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAAAACCGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAAAACCGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGCGGCTCTCACTGCTGTAAGAAAAAACCACCTCGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGCGGCTCTCACTGCTGTAAGAAAAAACCACCTCGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCACTGGCCACCAAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCACTGGCCACCAAGGTTT 1140
QY 1141 CCCGACTGAAAGCGGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGAAAGCGGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATATATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATATATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATATATGTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATATATGTGGA 1440
QY 1441 ATTGTGAGCGGATACAAATTTACACAGAAACAGCCAGTCGTTTAGTGTTTTTCAGA 1500
Db 1441 ATTGTGAGCGGATACAAATTTACACAGAAACAGCCAGTCGTTTAGTGTTTTTCAGA 1500
QY 1501 GCACCTTCAACAAAGGACCATAGATTATGAAATCGAAGAGGTAACTGCTAATCTCG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATTATGAAATCGAAGAGGTAACTGCTAATCTCG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTGGGCACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTGGGCACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCCGTAGCTTACACGCGAAGCTGATTCCTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCCGTAGCTTACACGCGAAGCTGATTCCTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGTCGCAACCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGTCGCAACCCGCAAAACCTGGGAA 1920
```

```
QY 1921 GAGATCCCGGCTGGATAAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCTCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCTCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGGCGTGGATTAACGCTGGCGCAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGGCGTGGATTAACGCTGGCGCAAGCG 2100
QY 2101 GGTCTGACCTTCTCGGTTGACCTGATTAAGAAACACATCAATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGGTTGACCTGATTAAGAAACACATCAATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGAAGAGCTGCGCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGAAGAGCTGCGCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCATCGACACAGCAAAAGTGAATTAATGTGTAAACGGTACTGCGCACTTC 2280
Db 2221 GCATGGTCCAAATCATCGACACAGCAAAAGTGAATTAATGTGTAAACGGTACTGCGCACTTC 2280
QY 2281 AAGGTCACACCATCCAAACCGTTGCTGGCGTCTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCATCCAAACCGTTGCTGGCGTCTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCAAGATCCACGTTATGCGCCCAACATGGAAGACCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCAAGATCCACGTTATGCGCCCAACATGGAAGACCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGTGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTCTGTGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACAAACAAATAACAAACCACTCGCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATAACAAACCACTCGCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTCTGTGGCCCGCGCGCTTCGAGCCCGG 2741
Db 2701 GGATCCTCTCTCTGTGGCCCGCGCGCTTCGAGGAGG 2741

RESULT 7
AX172306
LOCUS AX172306 8101 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 5 from Patent WO014467.
ACCESSION AX172306
VERSION AX172306.1 GI:14597487
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGNISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Goubin-Gramatica,F., Ducommun,B. and Prevost,G.
TITLE Method for obtaining human cdc25 phosphatases and method for identifying human cdc25 phosphatase modulators
JOURNAL Patent: WO 014467-A 5 21-JUN-2001; SOCIETE DE CONSEILS DE RECHERCHES ET D'APPLICATIONS SCIENTIFIQUES
```


FEATURES		(FR)		Location/Qualifiers	
source		1..8101		/organism="synthetic construct"	
				/mol_type="unassigned DNA"	
				/db_xref="taxon:32630"	
				/note="souche JM109 / pMAL - Hs Cdc25C"	
ORIGIN					
Query Match 82.0%; Score 2705.8; DB 6; Length 8101;					
Best Local Similarity 99.9%; Pred. No. 0;					
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA	60		
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA	60		
Qy	61	GTCAATTTCAGGGTGGTGAATGTGAACCAAGTAAGCTTATACGATGTGCGAGATATGCCG	120		
Db	61	GTCAATTTCAGGGTGGTGAATGTGAACCAAGTAAGCTTATACGATGTGCGAGATATGCCG	120		
Qy	121	GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACCGCCAGCCAGCTTTCTCGGAAA	180		
Db	121	GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACCGCCAGCCAGCTTTCTCGGAAA	180		
Qy	181	CGCGGGAAGAGTGGAAAGCGCGATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240		
Db	181	CGCGGGAAGAGTGGAAAGCGCGATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240		
Qy	241	AACAATCGCGGGCAACAGTCGTGTGATGTGGGTGGCCACCTCAGTCGTGCCCTGCG	300		
Db	241	AACAATCGCGGGCAACAGTCGTGTGATGTGGGTGGCCACCTCAGTCGTGCCCTGCG	300		
Qy	301	ACGCGCGTTCGCAAAATGTTCGCGCGATTAATCTCGCGCGATCAATCTGGGTGCCAGCG	360		
Db	301	ACGCGCGTTCGCAAAATGTTCGCGCGATTAATCTCGCGCGATCAATCTGGGTGCCAGCG	360		
Qy	361	TGTTGTGTGATGTGAACGAAGCGCGTCTGAAGCTGTGAAGCGCGGTGCACAATC	420		
Db	361	TGTTGTGTGATGTGAACGAAGCGCGTCTGAAGCTGTGAAGCGCGGTGCACAATC	420		
Qy	421	TTCTCGCGCAACCGTCAAGTGGGTGATCAATTAATCTCGCGCGATCAATCTGGGTGCCAG	480		
Db	421	TTCTCGCGCAACCGTCAAGTGGGTGATCAATTAATCTCGCGCGATCAATCTGGGTGCCAG	480		
Qy	481	TTGCTGTGGAGCTGCCTGCACTAATGTTCGCGCGTATTTCTTGATGTCTCTGACCAGA	540		
Db	481	TTGCTGTGGAGCTGCCTGCACTAATGTTCGCGCGTATTTCTTGATGTCTCTGACCAGA	540		
Qy	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600		
Db	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600		
Qy	601	TGTTGTGCAATTTGGGTCAACAGCAATTCGCTGTGAGCGGGCCCAATTAAGTTCTGTCG	660		
Db	601	TGTTGTGCAATTTGGGTCAACAGCAATTCGCTGTGAGCGGGCCCAATTAAGTTCTGTCG	660		
Qy	661	CGGCTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATG	720		
Db	661	CGGCTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATG	720		
Qy	721	CGGAAACGGGAGGCGATGGAGTGGCATATCTCACTCGCAATCAAAATTCAGCCGATG	780		
Db	721	CGGAAACGGGAGGCGATGGAGTGGCATATCTCACTCGCAATCAAAATTCAGCCGATG	780		
Qy	781	ATGAGGCGCATCTGTTCCCACTCGCATGTGGTGGCAACGATCAGATGGCGCTGGCGCAA	840		
Db	781	ATGAGGCGCATCTGTTCCCACTCGCATGTGGTGGCAACGATCAGATGGCGCTGGCGCAA	840		
Qy	841	TGCGCGCATTTACGAGTCCGGGTGCGGTGTGGTGGGATATCTCGGTAGTGGGATACG	900		
Db	841	TGCGCGCATTTACGAGTCCGGGTGCGGTGTGGTGGGATATCTCGGTAGTGGGATACG	900		

Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGTTAAACACCATCAAAACAGGATTTTC	960		
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGTTAAACACCATCAAAACAGGATTTTC	960		
Qy	961	GCTGTCTGGGGCAAAACAGCGTGGACCGCTTGTGTGCAACTCTCTCAGGCGCAGCGGTGA	1020		
Db	961	GCTGTCTGGGGCAAAACAGCGTGGACCGCTTGTGTGCAACTCTCTCAGGCGCAGCGGTGA	1020		
Qy	1021	AGGCAATCAGCTGTGGCGCTCTCACTGCTGAAGAAAACACACCTGGCGCCCAATA	1080		
Db	1021	AGGCAATCAGCTGTGGCGCTCTCACTGCTGAAGAAAACACACCTGGCGCCCAATA	1080		
Qy	1081	CGAAAACCGGCTCTCCCGCGGTTGGCGATTAATTAATGTCAGCTGGCAGCAGGTTT	1140		
Db	1081	CGAAAACCGGCTCTCCCGCGGTTGGCGATTAATTAATGTCAGCTGGCAGCAGGTTT	1140		
Qy	1141	CCGCACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200		
Db	1141	CCGCACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200		
Qy	1201	GCACAAATTCATGTTTGCAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260		
Db	1201	GCACAAATTCATGTTTGCAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260		
Qy	1261	TCAGGACGACCATCGGAAGCTGTGTATGGCTGTCCAGGTGCTAAATCACTGCATAATTCG	1320		
Db	1261	TCAGGACGACCATCGGAAGCTGTGTATGGCTGTCCAGGTGCTAAATCACTGCATAATTCG	1320		
Qy	1321	TGTGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCGCGCAGCATCAATAACGGTT	1380		
Db	1321	TGTGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCGCGCAGCATCAATAACGGTT	1380		
Qy	1381	CTGCGAAATATCTGAAATGAGTGTGTGACAAATTAATCATCGCTGCTATAATGTGTGGA	1440		
Db	1381	CTGCGAAATATCTGAAATGAGTGTGTGACAAATTAATCATCGCTGCTATAATGTGTGGA	1440		
Qy	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTTCACGA	1500		
Db	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTTCACGA	1500		
Qy	1501	GCACTTCAACCAACAGGACCATAGATTATGAAAATCTGAAGAAGTAAATCTGGTAAATCTGG	1560		
Db	1501	GCACTTCAACCAACAGGACCATAGATTATGAAAATCTGAAGAAGTAAATCTGGTAAATCTGG	1560		
Qy	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620		
Db	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620		
Qy	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCCGGCAAAAGCGTTCCAGGACAAAGCTGTAT	1680		
Db	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCCGGCAAAAGCGTTCCAGGACAAAGCTGTAT	1680		
Qy	1681	CGCGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACAGCCGCTTGGTGGCTAC	1740		
Db	1681	CGCGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACAGCCGCTTGGTGGCTAC	1740		
Qy	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGACAAAGCTGTAT	1800		
Db	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGACAAAGCTGTAT	1800		
Qy	1801	CGGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTCGCTTACCCGATCGCTGTT	1860		
Db	1801	CGGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTCGCTTACCCGATCGCTGTT	1860		
Qy	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCCAAACCTGGGAA	1920		
Db	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCCAAACCTGGGAA	1920		
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGTGTTCAAC	1980		
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGTGTTCAAC	1980		
Qy	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG	2040		

Db 1981 CTGCAAGAACCGTACTTACCTGGCGCGTGATTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAAGTGGCGGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAAGTGGCGGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAATAAACAACACATGAATGCGAGACACCATTAC 2160
Db 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAATAAACAACACATGAATGCGAGACACCATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTCTTAATAAAGCGGAAAAAGAGATGACCATCAAGCGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTCTTAATAAAGCGGAAAAAGAGATGACCATCAAGCGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACACAGCAAGATGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACAGCAAGATGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCAACCATCAAAACCGTTCTTGGCGTGTGAGGCGAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTCTTGGCGTGTGAGGCGAGGTATTAAACGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGCGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGCGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCACGTATTGCGCGCCACATGGAAGAACGCCAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTGCGCGCCACATGGAAGAACGCCAGAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCGTACGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCGTACGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGAGCGCGACACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGAGCGCGACACTAAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
AX377531
LOCUS AX377531 6648 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 8 from Patent WO0212553.
ACCESSION AX377531
VERSION AX377531.1 GI:19573717
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Kappel, A., Polakowski, T., Pignot, M., Windhab, N., Behrens, H. and Muth, J.
TITLE Method for detecting mutations in nucleotide sequences
JOURNAL Patent: WO 0212553-A 8 14-FEB-2002;
Nanogen Recognomics GmbH (DE)
FEATURES
Location/Qualifiers
1..6648
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 81.8%; Score 2701; DB 6; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
/note="Plasmid pMAL-c2x"
Qy 1 CCGACACCATGAAATGGTGGAAAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
Db 1 CCGACACCATGAAATGGTGGAAAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
Qy 61 GTCAATTTCAGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTCGACAGATATGCG 120
Db 61 GTCAATTTCAGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTCGACAGATATGCG 120
Qy 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCCAGCCACGTTTCTGCAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCCAGCCACGTTTCTGCAAAA 180
Qy 181 CCGGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTAATTCCTCCAAACCGCTGGC 240
Db 181 CCGGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTAATTCCTCCAAACCGCTGGC 240
Qy 241 AACAACTGCGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGGCCCTGCGCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCGCAGCG 360
Db 301 ACGGCCCTGCGCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCGCAGCG 360
Qy 361 TGGTGGTTCGATGTTGTAAGAAAGAGCGCGCTGAAAGCCCTGTAAGCGCGGTGCAATTC 420
Db 361 TGGTGGTTCGATGTTGTAAGAAAGAGCGCGCTGAAAGCCCTGTAAGCGCGGTGCAATTC 420
Qy 421 TTCTCGCGCAACCGCTGAGTGGCGTGCATTAATCTATCTCGTGAATGACACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTGAGTGGCGTGCATTAATCTATCTCGTGAATGACACAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGCTTATTTCTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGCTTATTTCTGATGCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGCTCGG 660
Db 601 TGGTGGCATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CCGAAACGGGAAGCGGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAAACGGGAAGCGGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATAGGGCATGTTCCCACTGGGATGCTGTTGCCAAGCATGATGGCGCTGGGGCGCA 840
Db 781 ATAGGGCATGTTCCCACTGGGATGCTGTTGCCAAGCATGATGGCGCTGGGGCGCA 840
Qy 841 TGGCGCCATTTACCGAGTCCGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGCCATTTACCGAGTCCGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCTCAACTCTCTCAGGGCGAGGGTGA 1020
Db 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCTCAACTCTCTCAGGGCGAGGGTGA 1020

Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCACACTCTCTCAGGCGCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGGCCGCTCTCAGCTGGTGAAGAAAAACACACCTCGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCCGCTCTCAGCTGGTGAAGAAAAACACACCTCGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTGGCCGATTCATTAAATGAGTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTGGCCGATTCATTAAATGAGTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAAGCAACCAATTAATGTAGTGTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCAACCAATTAATGTAGTGTAGCTCACTCATTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGTATGGTGTGCAAGTCTGAATCACTGCAATAATTTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGTATGGTGTGCAAGTCTGAATCACTGCAATAATTTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGATAATGTTTTCGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGATAATGTTTTCGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAACAGCGAGTCCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAACAGCGAGTCCGTTTAGGTGTTTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAACACTGAAGAGTAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAACACTGAAGAGTAACTGGTAATCTGG 1560
Qy 1561 ATTAACCGCGATAAAGGCTTAACCGGTCTCGCTGAAGTTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTTAACCGGTCTCGCTGAAGTTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGCGATGCGCTGCATTAATCTCTGGGCAACAGCGCTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGATGCGCTGCATTAATCTCTGGGCAACAGCGCTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAAATCACCCCGACAAAGCGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCACCCCGACAAAGCGTTCCAGGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTACGTTACAAAGGCTGATGTTTACCGGATGCGTGT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTACAAAGGCTGATGTTTACCGGATGCGTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCGAACCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCGAACCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACCGCAAGTACGACATTAAGACATGCGGCGTGGATACGCTGCGGAAAGCG 2100
Db 2041 TATGAAAAACCGCAAGTACGACATTAAGACATGCGGCGTGGATACGCTGCGGAAAGCG 2100

Qy 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAAAAACAAACACATGAATCGAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAAAAACAAACACATGAATCGAGACCGGATTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACCGGCCCTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACCGGCCCTGG 2220
Qy 2221 GCATGTCCAAATCGACACACAGCAAGTGAATTAATGTTGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGTCCAAATCGACACACAGCAAGTGAATTAATGTTGTAAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCAATCCAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACCGGCCAGT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACCGGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGTTTAAATAAGCAAAACCGTGGTGGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGCAAAACCGTGGTGGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATCGAAACCGCCAGAAAGCTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATCGAAACCGCCAGAAAGCTGAAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGCGTCTTCTGTTGTCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGAGATGTCGCGTCTTCTGTTGTCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGAGACTGTCGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGAGACTGTCGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAAC 2700
Db 2641 AACAAACAAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAAC 2700
Qy 2701 GGATCTCT 2709
Db 2701 GGATCTCT 2709

RESULT 9
A91965
LOCUS 7475 bp DNA circular PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9822496.
ACCESSION A91965
VERSION A91965.1 GI:6740811
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood, M. R. and Hurst, D. N.
TITLE ANTIVIRAL PEPTIDE DERIVATIVES
JOURNAL Patent: WO 9822496-A 1 28-MAY-1998;
HOFFMANN LA ROCHE (CH)
FEATURES
source 1. .7475
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATCGTGCAAAACCTTTTCGGGTATGGCATGATGATGCGGCCCGGAAAGAGA 60

|||||
1 CCGACCAATCAATGGGTGCAAAACCTTTGCGGGTATGGCATGTAGCGCCCGGAAGAGA 60
61 GTCAATTGAGGGTGTGAATGTGAACACCAAGTAACTGATGTCGACAGTATGCGG 120
61 GTCAATTGAGGGTGTGAATGTGAACACCAAGTAACTGATGTCGACAGTATGCGG 120
121 GTGTCCTTATCAGACCGTTTCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAAA 180
121 GTGTCCTTATCAGACCGTTTCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAAA 180
181 CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTAACATCCCAACCGGTGGCAC 240
181 CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTAACATCCCAACCGGTGGCAC 240
241 AACAACTGGCGGCAAAACAGTGGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
241 AACAACTGGCGGCAAAACAGTGGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
301 AGCGCCGTGCGAAATTTGTGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360
301 AGCGCCGTGCGAAATTTGTGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360
361 TGGTGGTGTGATGCTAGAACGAAAGCGCGTGGAGCGCTGTAAGCGCGGTGCAAAATC 420
361 TGGTGGTGTGATGCTAGAACGAAAGCGCGTGGAGCGCTGTAAGCGCGGTGCAAAATC 420
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCGCGTATTTCTGATGTCTGACCCAGA 480
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCGCGTATTTCTGATGTCTGACCCAGA 480
481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTATTTCTGATGTCTCTGACCCAGA 540
481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTATTTCTGATGTCTCTGACCCAGA 540
541 CACCCATCAACAGTATTTCTCCCATGAAGCGGTACCGCATGCGCGTGGAGCATC 600
541 CACCCATCAACAGTATTTCTCCCATGAAGCGGTACCGCATGCGCGTGGAGCATC 600
601 TGGTGGCATTTGGTCAACAGCAAACTGGCGTGTAGCGGGCCCAATTAAGTTCGTCTCGG 660
601 TGGTGGCATTTGGTCAACAGCAAACTGGCGTGTAGCGGGCCCAATTAAGTTCGTCTCGG 660
661 CGCGTCTGCTGTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
661 CGCGTCTGCTGTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCA 840
781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCA 840
841 TCGCGCCATTAACGAGTCCGGCTGCGCGTTGGTGGGATATCTCGTGGGATAG 900
841 TCGCGCCATTAACGAGTCCGGCTGCGCGTTGGTGGGATATCTCGTGGGATAG 900
901 ACGATACCGAGACAGCTCATGTTATCCCGCGTAAACACCATCAACAGGATTTTC 960
901 ACGATACCGAGACAGCTCATGTTATCCCGCGTAAACACCATCAACAGGATTTTC 960
961 GCCTGTGGGCAAAACACGCTGGACCGTCTGCTGCACTCTCTCAGGGCCAGGCGGTGA 1020
961 GCCTGTGGGCAAAACACGCTGGACCGTCTGCTGCACTCTCTCAGGGCCAGGCGGTGA 1020
1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGCAAGCTGGCAGCAGAGTTT 1140
|||||

Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGCAAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGAACAGCTTATCATCGACTGCAAGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAACAGCTTATCATCGACTGCAAGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTGTGCTGTGAGGTGCGAGGTGCTAATCACTGCAATTTTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTGTGCTGTGAGGTGCGAGGTGCTAATCACTGCAATTTTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTCTTGGATAATGTTTTTTCGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTCTTGGATAATGTTTTTTCGCGGACATCAATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
QY 1441 ATTGTAGCGGATTAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAGCA 1500
Db 1441 ATTGTAGCGGATTAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAGCA 1500
QY 1501 GCATTTACCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCATTTACCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTTAACCGCGATAAGGCTATAACGGTCTCCGCTGAGTCCGTTAGAGTTCGAGAAAGAT 1620
Db 1561 ATTTAACCGCGATAAGGCTATAACGGTCTCCGCTGAGTCCGTTAGAGTTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
QY 1681 CGGCAACTGCGATGCGCTTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGATGCGCTTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGCAAGCTGATTTGCTTACCGATCGCTGT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGCAAGCTGATTTGCTTACCGATCGCTGT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAAACCGTACTTCCACTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAAACCGTACTTCCACTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAAACGTTGGCGTGGATTAACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAAACGTTGGCGTGGATTAACGCTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAAACATGATGAGACACCGATTTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAAACATGATGAGACACCGATTTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAAGCGGCCGTTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAAGCGGCCGTTGG 2220

QY	1321	TGTCGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTTGGCGCGACATCATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCGACCTCCCGTTCTGGATAATGTTTTTTGGCGCGACATCATACCGTT	1380
QY	1381	CTGGCAAAATATTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Db	1381	CTGGCAAAATATTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
QY	1441	ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCGGTTTAGTGTTTTTACGA	1500
Db	1441	ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCGGTTTAGTGTTTTTACGA	1500
QY	1501	GCACCTTCAACACAGGACCATAGATTATGAAAACCTGAAGAAGTTAAACTGGTAAATCTGG	1560
Db	1501	GCACCTTCAACACAGGACCATAGATTATGAAAACCTGAAGAAGTTAAACTGGTAAATCTGG	1560
QY	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
QY	1681	GGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	GGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCTCTGGCTGAATCAACCCCGGAAGCGTTCCAGGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCTCTGGCTGAATCAACCCCGGAAGCGTTCCAGGACAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGCGCTAGCTTACAAACGGCAAGCTGATGCTTACCCGATCGCTGTT	1860
Db	1801	CGGTTTACCTGGGATGCGCTAGCTTACAAACGGCAAGCTGATGCTTACCCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCGCAAAAACTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCGCAAAAACTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGTGACGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGTGACGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGATAACCGCTGGCGGAAAGCG	2100
Db	2041	TATGAAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGATAACCGCTGGCGGAAAGCG	2100
QY	2101	GGTCTGACCTTCTGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC	2160
QY	2161	TCATCGCAGAAGCTGCCCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGCTGG	2220
Db	2161	TCATCGCAGAAGCTGCCCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGCTGG	2220
QY	2221	GCATGGTCCAAATCGACACCAAGCAAGTAAATATGTTGTTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGGTCCAAATCGACACCAAGCAAGTAAATATGTTGTTAAACGGTACTGCCGACCTTC	2280
QY	2281	AAGGGTCAACATCCAAACCGTTGTTGGCGTCTGAGCGCAGGTATTAAACCGCGCCAGT	2340
Db	2281	AAGGGTCAACATCCAAACCGTTGTTGGCGTCTGAGCGCAGGTATTAAACCGCGCCAGT	2340
QY	2341	CCGAACAAAGAGCTGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGGCTCTG	2400
Db	2341	CCGAACAAAGAGCTGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGGCTCTG	2400

QY	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG	2460
QY	2461	TTGGCGAAAGATCCACGCTATTGGCCGCCACCATGGAAAAACGCCAGAAAAGGTGAAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGCTATTGGCCGCCACCATGGAAAAACGCCAGAAAAGGTGAAATCATG	2520
QY	2521	CGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCGGTGATCAACGCC	2580
Db	2521	CGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCGGTGATCAACGCC	2580
QY	2581	GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
QY	2641	AACAAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2641	AACAAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700

RESULT 11

AR207294

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..7475

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match

Best Local Similarity

Matches 2700;

Conservative 0;

Mismatches 0;

Indels 0;

Gaps 0;

QY

1

CGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGCGATGATAGCGCCCGGAAGAGA

60

Db

1

CGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGCGATGATAGCGCCCGGAAGAGA

60

QY

61

GTCAATTCAGGGTGGTGAATGTGAAACCAAGTAACGTTATACGATGTCGAGAGTATGCCG

120

Db

61

GTCAATTCAGGGTGGTGAATGTGAAACCAAGTAACGTTATACGATGTCGAGAGTATGCCG

120

QY

121

GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAAACAGGCCAGCCACGTTTCTGGCAAAA

180

Db

121

GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAAACAGGCCAGCCACGTTTCTGGCAAAA

180

QY

181

CGCGGAAAAGTGGNAGCGCGATGGGAGCTGAATTAATCCCAACCGGTGGCAC

240

Db

181

CGCGGAAAAGTGGNAGCGCGATGGGAGCTGAATTAATCCCAACCGGTGGCAC

240

QY

241

AACAACTGGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC

300

Db

241

AACAACTGGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC

300

QY

301

ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG

360

Db

301

ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG

360

QY

361

TGGTGGTCTCGATGTAAGCAAGCGCGCTCGAAGCCCTGTAAGCGCGCGGTGCACATC

420

Db

361

TGGTGGTCTCGATGTAAGCAAGCGCGCTCGAAGCCCTGTAAGCGCGCGGTGCACATC

420

QY 421 TTCTCGCGAAGCGCTCAGTGGCTGATCAATTAACATATCCGTGGATGACCAAGATGCCA 480
DB |||||
QY 421 TTCTCGCGAAGCGCTCAGTGGCTGATCAATTAACATATCCGTGGATGACCAAGATGCCA 480
DB |||||
QY 481 TTGCTGTGGAAAGCTCGCTGCACATTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCAAGA 540
DB |||||
QY 481 TTGCTGTGGAAAGCTCGCTGCACATTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCAAGA 540
DB |||||
QY 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 600
DB |||||
QY 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 600
DB |||||
QY 601 TGGTCGCAATTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTGG 660
DB |||||
QY 601 TGGTCGCAATTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTGG 660
DB |||||
QY 661 CGCGTCTCGCTCTGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB |||||
QY 661 CGCGTCTCGCTCTGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB |||||
QY 721 CGGAAACGGAAAGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB |||||
QY 721 CGGAAACGGAAAGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB |||||
QY 781 ATGAGGGCATGTTCCCATCGCATGCTGGTGGCAACGATCGATGGCGTGGGCGCAA 840
DB |||||
QY 781 ATGAGGGCATGTTCCCATCGCATGCTGGTGGCAACGATCGATGGCGTGGGCGCAA 840
DB |||||
QY 841 TGGCGCCATTAACGAGTCCGGCTGGCGTGGTGGGATATCTCGTACTGGGATACG 900
DB |||||
QY 841 TGGCGCCATTAACGAGTCCGGCTGGCGTGGTGGGATATCTCGTACTGGGATACG 900
DB |||||
QY 901 ACGATACCGAAAGACGCTCATGTTATATCCCGCGTTAAACCAACATCAAAACAGGATTTTC 960
DB |||||
QY 901 ACGATACCGAAAGACGCTCATGTTATATCCCGCGTTAAACCAACATCAAAACAGGATTTTC 960
DB |||||
QY 961 GCCTGTGGGGCAAAACGAGCTGAGACCGCTTGTGTGCAACTCTTCAGGGCCAGCGGTGA 1020
DB |||||
QY 961 GCCTGTGGGGCAAAACGAGCTGAGACCGCTTGTGTGCAACTCTTCAGGGCCAGCGGTGA 1020
DB |||||
QY 1021 AGGCAATACGCTGTGCGCGTCTCACTGTGGTGAAGAAACCAACCACTGGCGCCCAATA 1080
DB |||||
QY 1021 AGGCAATACGCTGTGCGCGTCTCACTGTGGTGAAGAAACCAACCACTGGCGCCCAATA 1080
DB |||||
QY 1081 CGCAAAACGCGCTCTCCCGCGCTTGGCCGATTCAATTAATGCAGCTGGCACGACAGTTT 1140
DB |||||
QY 1141 CCGACTGGAAAGCGGCGAGTGAGGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
DB |||||
QY 1141 CCGACTGGAAAGCGGCGAGTGAGGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
DB |||||
QY 1201 GCACAAATCTCATGTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGCG 1260
DB |||||
QY 1201 GCACAAATCTCATGTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGCG 1260
DB |||||
QY 1261 TCAGGCGACCATCGGAAGCTGTGATGGCTGTGCGATTAATCACTGCATAATTCG 1320
DB |||||
QY 1321 TGTGCTCAAGGCGCATCTCCGTTCTGGATTAATGTTTTTGGCGCGGACATCAATACCGTT 1380
DB |||||
QY 1321 TGTGCTCAAGGCGCATCTCCGTTCTGGATTAATGTTTTTGGCGCGGACATCAATACCGTT 1380
DB |||||
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGGGA 1440
DB |||||
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGGGA 1440
DB |||||
QY 1441 ATTGTAGCGGATAAACATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
DB |||||
QY 1441 ATTGTAGCGGATAAACATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
DB |||||

QY 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGTAAACTGGTAACTCGG 1560
DB |||||
QY 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGTAAACTGGTAACTCGG 1560
DB |||||
QY 1561 ATTAACCGCGATTAAGGCTATTAACGGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1561 ATTAACCGCGATTAAGGCTATTAACGGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1621 ACCGGAATTAAGTACCGCTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGTT 1680
DB |||||
QY 1621 ACCGGAATTAAGTACCGCTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGTT 1680
DB |||||
QY 1681 GCGCAACTGGCGATGCCCTGACATTTCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB |||||
QY 1681 GCGCAACTGGCGATGCCCTGACATTTCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB |||||
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGACAAAGCGTTCAGGACAACTGTAT 1800
DB |||||
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGACAAAGCGTTCAGGACAACTGTAT 1800
DB |||||
QY 1801 CCGTTTACCTGGGATGCCGTACGTTCAACCGCAAGCTGATTGCTTACCCTGATCGCTGTT 1860
DB |||||
QY 1801 CCGTTTACCTGGGATGCCGTACGTTCAACCGCAAGCTGATTGCTTACCCTGATCGCTGTT 1860
DB |||||
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGGAA 1920
DB |||||
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGGAA 1920
DB |||||
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1981 CTGCAAGAACCGTACTTCACTCGCGCTGATTGCTGACCGGGGTATGCGTTCAG 2040
DB |||||
QY 1981 CTGCAAGAACCGTACTTCACTCGCGCTGATTGCTGACCGGGGTATGCGTTCAG 2040
DB |||||
QY 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAATAACAAACATGATGACAGACCGGATAC 2160
DB |||||
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAATAACAAACATGATGACAGACCGGATAC 2160
DB |||||
QY 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACCGGCCGTGG 2220
DB |||||
QY 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACCGGCCGTGG 2220
DB |||||
QY 2221 GCATGTCGAACATCGACACAGCAAAAGTGAATTAAGTGTAAACGGTACTGCGGACCTTC 2280
DB |||||
QY 2221 GCATGTCGAACATCGACACAGCAAAAGTGAATTAAGTGTAAACGGTACTGCGGACCTTC 2280
DB |||||
QY 2281 AAGGTCACACATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT 2340
DB |||||
QY 2281 AAGGTCACACATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT 2340
DB |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
DB |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
DB |||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
DB |||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
DB |||||
QY 2461 TTGCGGAAAGATCCAGTATTTGCGCCACCATCGAAAAACCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2461 TTGCGGAAAGATCCAGTATTTGCGCCACCATCGAAAAACCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2521 CCGAAACATCCCGCAGATGTCGCTTTCTGGTATCCGTGCGTACTGCGGTGATCAACGCC 2580
DB |||||
QY 2521 CCGAAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
DB |||||
QY 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
DB |||||

Db 2581 |||||GCCAGCGTGCAGACTGTCGATGAAGCCCTGGAAGACGCGCAGCTAAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAACAATAACAATAACAACCTCGGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAACAATAACAATAACAACCTCGGGATCGAGGAGGATTTTCAGAAATTC 2700
RESULT 12
AX377532
LOCUS AX377532 9191 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 9 from Patent WO0212553.
ACCESSION AX377532
VERSION AX377532.1 GI:19573718
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdoerf,H. and
Muth,J.
TITLE Method for detecting mutations in nucleotide sequences
JOURNAL Patent: WO 0212553-A 9 14-FEB-2002;
Nanogen Recognition GmbH (DE)
FEATURES
source
1..9191
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pMALc2x-mutS Plaemid"
ORIGIN
Query Match 81.8%; Score 2698; DB 6; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGCGGTATGGCATATAGAGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGCGGTATGGCATATAGAGCCCGGAAGAGA 60
Qy 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACCATGTCCGAGAGTATGCG 120
Db 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACCATGTCCGAGAGTATGCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCCAGCCACGTTTCTGGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCCAGCCACGTTTCTGGGAAAA 180
Qy 181 CCGGGGAAAAAGTGAAGCGCGCATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CCGGGGAAAAAGTGAAGCGCGCATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGCGGGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGTGTCGATGTGTAGAACGAAGCGCGTGCAGACCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGGTGTGTCGATGTGTAGAACGAAGCGCGTGCAGACCTGTAAAGCGGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCAGA 540
Qy 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

Db 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTGTCTGTCTCGG 660
Db 601 TGGTGCATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTGTCTGTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCAATTCGAAATTCGTA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCAATTCGAAATTCGTA 780
Qy 781 ATGAGGGCATCGTTCCCACTCGCATGTGTTGCCAACGATCGATGCGCTGGCGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGCATGTGTTGCCAACGATCGATGCGCTGGCGCGCAA 840
Qy 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGCTGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCGCTCTCAGTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCAGTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGTGGCGGATTCATTAAATGAGCTGGCACACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGTGGCGGATTCATTAAATGAGCTGGCACACAGGTTT 1140
Qy 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTTAG 1200
Db 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTTAG 1200
Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGTTCTGGCG 1260
Qy 1261 TCAGGACCGCATCGGAAGCTGTGTATGGCTGTGAGCTCGTAAATCATCTGCNTAATTCG 1320
Db 1261 TCAGGACCGCATCGGAAGCTGTGTATGGCTGTGAGCTCGTAAATCATCTGCNTAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATAACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Qy 1441 ATTGTAGCGGATTAACAATTTTACACAGGAACACGCGCTCGGTTAGTGTTCACCA 1500
Db 1441 ATTGTAGCGGATTAACAATTTTACACAGGAACACGCGCTCGGTTAGTGTTCACCA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATATGAANAATCGAAGAGGTTAACTGGTAAATCTCG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATATGAANAATCGAAGAGGTTAACTGGTAAATCTCG 1560
Qy 1561 ATTTAAGCGGATAAAGGCTATAACCGGCTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTATAACCGGCTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAAATCCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATATCTCTCGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATATCTCTCGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GGTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GGTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Qy 1801 CGTTTACTCGGATGCGGTAGCTTACACCGCAAGCTGATGCTTACCGATGCTGTT 1860
Db 1801 CGTTTACTCGGATGCGGTAGCTTACACCGCAAGCTGATGCTTACCGATGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAGAACTGAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAGAACTGAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACGCTACTCTCACCTGGCGCTGATGCTGCTGACGGGGTTATCGCTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTCTCACCTGGCGCTGATGCTGCTGACGGGGTTATCGCTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTGGTTGACCTGATTAAGAAACAAACACATGAATGACGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAAGAAACAAACACATGAATGACGACCGGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACAGCGATGACCATCAACGCGCCGCTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACGACCAAGTGAATTAAGTGTAAACGGTACTCGGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACGACCAAGTGAATTAAGTGTAAACGGTACTCGGACCTTC 2280
Qy 2281 AAGGTCACCATCCAAACCGTTCTGGCGTGTGACGCGCAGGTATTAACGCGCCGCACT 2340
Db 2281 AAGGTCACCATCCAAACCGTTCTGGCGTGTGACGCGCAGGTATTAACGCGCCGCACT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACCTATCTGCTGACTGATGAGGTTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACCTATCTGCTGACTGATGAGGTTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCCGCTGGTCCGTTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCCGCTGGTCCGTTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGCGGAAAGATCCACGTTATGCGCCACCATGAAACCGCCGCAAGAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTTATGCGCCACCATGAAACCGCCGCAAGAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCGAGCGGTGCTGAGCTGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCGAGCGGTGCTGAGCTGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCTCGGATCGAGGAGGATTCAGATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGATCGAGGAGGATTCAGATTC 2700
Qy 2701 GGATCC 2706
Db 2701 GGATCC 2706

RESULT 13
AF097412
LOCUS
DEFINITION
AF097412
ACCESSION
AF097412.1 GI:3983120
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 6724)
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 6724)
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..6724
/organism="Expression vector pMBP-parallel"
/mol_type="other DNA"
/db_xref="taxon:84479"
204..1163
/gene="lacIq"
204..1163
/gene="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAC3812.1"
/db_xref="GI:3983121"
/translations="MAELNYIPNRVAQLAGKQSLIGVATSLAHAPSQIVAAIKS
RADQLGASVVSVMSVERSGVEACKAAVHNLQAQRVGLIINPLDDQDAIAVEACTNV
PALFLDVSQTPINSIIFSHEDGTRGLVGHVVALGHQIALLAGPLSSVARSRLAGW
HKYLTNRQIPIAEREGDWSAMSGFOQTQMLNEGIPTAMLVANDQMALGAMRAITE
SGLRVGADISVVGDDTDESSCYIPPLTTIKODFRLLGOTSDVRLILQLSQGQAVKGNQ
LPLVSLVKRKTTLAPNTQTASPRALADSLMQLARQVSRLESG"
1453..1481
/notes="tac; inducible promoter"
1528..2994
/genes="male"
1528..2994
/genes="male"
/codon_start=1
/transl_table=11
/product="maltose binding protein"
/protein_id="AAC3813.1"
/db_xref="GI:3983122"
/translations="MKIERGLVITWINGDKYNGLAEVGKPEKDTGKIVTVEHPDKL
EKFPVAATGDPDIIFWAHDPRPGYAGSGLLAETPKAFQDKLYLPFTWDVRYNG
KLIAYIAVEALSLIYNKOLLNPPTKWEIIPALDKELKAKGSAFLMFLNQEPYFTWP
LIAADGGYAPKPVNGKIDYKVDNAGAKAGTLFLVDLTKNKHMDYDYSABAA
NGETAMTTNGPWAMSNIDTSKVNYGTVTLPTFKGQSPKFPVGLSAGINAA
AKEFLENYLLTDEGLEAVNKKPLGVALKSYEEELAKDPRIATAATMENAKGIMPN
PQMSAFWAVRTAVINAAAGROTVDALKAQDNTSSNNNNNNNNNNNNLIGDTTENVLP
QAMDPKPLRRRAQLVRPLNLEPAVSRHAPVPSIALAVLQRRDWNENPGVTQLNRL
AAHPFPASWENSEARTDRPSQQLRSNLGWLGCFFG"
2633..2688
/genes="male"
/notes="protein spacer region"
2689..2709
/genes="male"
/notes="Tobacco etch virus protease cleavage site"
2711..2810
/genes="male"

gene	/note="multiple cloning site"									
	3426..4356									
	/gene="bla"									
promoter	3426..3454									
	/gene="bla"									
	3496..4356									
	/gene="bla"									
	/codon_start=1									
	/transl_table=11									
	/product="beta-lactamase"									
	/protein_id="AAC83814.1"									
	/db_xref="GI:3983123"									
	/translation="MSIQHFRVALIIPFAAFLPVFAHPETLVKVKDAEDQLGARVGY									
	IELDNSKRIEFSPEERFSFMSTFKVLLCGAVLSRVDAQGEQQLGRIRHYSDNDLVE									
	YSPVTEKHLDGMTVELCSAAITMSDNTAANLLLTITGGPKELTAPLHNMGDHTVRL									
	DRWPELNEAIPNDRDITMPVMAATIKLLTIGELLTLASRQQLIDNWEADKVGAPL									
	LRSALPAGWFIADKSGAGBSRGI I AALGPDGKPSRIWVITYTGSQATMDERNQIA									
	EIGASLIKHW"									
ORIGIN										
Query Match	81.1%; Score 2676.2; DB 12; Length 6724;									
Best Local Similarity	99.9%; Pred. No. 0;									
Matches 2678; Conservative	0; Mismatches 3; Indels 0; Gaps 0;									
QY	1	CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA	60							
DB	1	CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA	60							
QY	61	GTCAATTACGGGTGGTGAATGTGAACACGTAACGTTATACGATGTCGCAGAGTATGCCG	120							
DB	61	GTCAATTACGGGTGGTGAATGTGAACACGTAACGTTATACGATGTCGCAGAGTATGCCG	120							
QY	121	GTGTCTCTTATCAGACACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA	180							
DB	121	GTGTCTCTTATCAGACACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA	180							
QY	181	CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTACATTCGCCAACCGGTGGCAC	240							
DB	181	CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTACATTCGCCAACCGGTGGCAC	240							
QY	241	AACAACTGCGGGCAACAGTCGTTGCTGATTGGCGTGTGCACCTCCAGTCTGGCCCTGC	300							
DB	241	AACAACTGCGGGCAACAGTCGTTGCTGATTGGCGTGTGCACCTCCAGTCTGGCCCTGC	300							
QY	301	ACGCGCGTTCGCAAAATGTTCGCGCGCATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360							
DB	301	ACGCGCGTTCGCAAAATGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360							
QY	361	TGGTGTGTGATGTAGAACGAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC	420							
DB	361	TGGTGTGTGATGTAGAACGAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC	420							
QY	421	TTCTCGCGCAACGGTCAAGTGGGCTGATCATTAATCTCGCTGGATGACCAAGATGCCA	480							
DB	421	TTCTCGCGCAACGGTCAAGTGGGCTGATCATTAATCTCGCTGGATGACCAAGATGCCA	480							
QY	481	TTGCTGTGGAAGCTGCTGCACATAATGTTCGCGCGTTATTTCTTGATGTCTCTGACCAGA	540							
DB	481	TTGCTGTGGAAGCTGCTGCACATAATGTTCGCGCGTTATTTCTTGATGTCTCTGACCAGA	540							
QY	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600							
DB	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600							
QY	601	TGGTCCGATTTGGGTTCACCGAAATCGCGCTGTAGCGGGCCATTAAGTTCTGTCTCGG	660							
DB	601	TGGTCCGATTTGGGTTCACCGAAATCGCGCTGTAGCGGGCCATTAAGTTCTGTCTCGG	660							
QY	661	CGCGTCTGCGTCTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720							
DB	661	CGCGTCTGCGTCTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720							
QY	721	CGGAACGGGAAGGCGAATGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGTA	780							

DB	721	CGGAACGGGAAGCGGACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA	780	
QY	781	ATGAGGGCATCGTTCCTCCACTCGCATGTGGTGGCAACGATCAGATGGCGTGGCGCAA	840	
DB	781	ATGAGGGCATCGTTCCTCCACTCGCATGTGGTGGCAACGATCAGATGGCGTGGCGCAA	840	
QY	841	TGCGGCCATTACCGAGTCCGGGTTCGGGTGGTGGGATATCTCGGTAGTGGGATAG	900	
DB	841	TGCGGCCATTACCGAGTCCGGGTTCGGGTGGTGGGATATCTCGGTAGTGGGATAG	900	
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACACCATCAAAACAGGATTTTC	960	
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACACCATCAAAACAGGATTTTC	960	
QY	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGGCGTGA	1020	
DB	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGGCGTGA	1020	
QY	1021	AGGGCAATCAGCTGTGGCCCTCTCAGCTGTGAAAGAAAAACCACTGGCGGCCAATA	1080	
DB	1021	AGGGCAATCAGCTGTGGCCCTCTCAGCTGTGAAAGAAAAACCACTGGCGGCCAATA	1080	
QY	1081	CGCAACCCCTCTCCCGCGCTTGGCCGATTCAATTAATGCAGCTGCACGACAGGTTT	1140	
DB	1081	CGCAACCCCTCTCTCCCGCGCTTGGCCGATTCAATTAATGCAGCTGCACGACAGGTTT	1140	
QY	1141	CCGACTCGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200	
DB	1141	CCGACTCGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200	
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGTTCTGGCG	1260	
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGTTCTGGCG	1260	
QY	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATTAATTCG	1320	
DB	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATTAATTCG	1320	
QY	1321	TGTCCTCAAGCGCACTCCCGCTTCGGATAATGTTTTTGGCGCGACATCAATAACGGTT	1380	
DB	1321	TGTCCTCAAGCGCACTCCCGCTTCGGATAATGTTTTTGGCGCGACATCAATAACGGTT	1380	
QY	1381	CTGGCAAAATTTCTGAAATGAGCTTTCACAGCAATTAATCAATCGGCTCGTATATGTGTGA	1440	
DB	1381	CTGGCAAAATTTCTGAAATGAGCTTTCACAGCAATTAATCAATCGGCTCGTATATGTGTGA	1440	
QY	1441	ATTGTGAGCGGATAACAAATTTTCACAGCAACAGCCAGTCCGTTTAGTGTTCACGA	1500	
DB	1441	ATTGTGAGCGGATAACAAATTTTCACAGCAACAGCCAGTCCGTTTAGTGTTCACGA	1500	
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG	1560	
DB	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG	1560	
QY	1561	ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT	1620	
DB	1561	ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT	1620	
QY	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680	
DB	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680	
QY	1681	GGGGCAATGGCGATGGCCCTGACATTTCTTCTGGGACACGCGCTTTGGTGGCTAC	1740	
DB	1681	GGGGCAATGGCGATGGCCCTGACATTTCTTCTGGGACACGCGCTTTGGTGGCTAC	1740	
QY	1741	GCTCAATCTGGCTGTGGCTGAAAATCAACCCCGGCAAAAGCGTTCAGGACAAGCTGTAT	1800	
DB	1741	GCTCAATCTGGCTGTGGCTGAAAATCAACCCCGGCAAAAGCGTTCAGGACAAGCTGTAT	1800	
QY	1801	CGGTTTACCTGGGATGCGGTACGTTTACACGGCAAGCTGATTGCTTACCCGATCGCTGTT	1860	

Db 1801 CCGTTTAACTGGGATCCGATAGCTTAAACGCGAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTCCGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTCGATGAAGAACTGAAGCAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATGAAGAACTGAAGCAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGTTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGTTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGAAGCTGGCGCTGGATAAGCTGGCGGCGAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGAAGCTGGCGCTGGATAAGCTGGCGGCGAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGCGGATGATGATGATGATGATGATGATGATG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGCGGATGATGATGATGATGATGATGATGATG 2220
QY 2221 GCATGGTCCAAATCGACACACGACGAAAGTGAATATGTTGTAACGCTGCTCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGACGAAAGTGAATATGTTGTAACGCTGCTCCGACCTTC 2280
QY 2281 AAGGTCAACCATCAAAACCGTTCTGTTGGCTGCTGAGCGAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTCTGTTGGCTGCTGAGCGAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCTGCTGAGCTGAAAGTCTTACGAGAAAG 2460
Db 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCTGCTGAGCTGAAAGTCTTACGAGAAAG 2460
QY 2461 TTGGGAAAGATCCACGTATTTGCCGCCACATGAAAGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTATTTGCCGCCACATGAAAGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGCTCGCTTTCTGTTATGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGCTCGCTTTCTGTTATGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTCTGACATGCTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGCTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATTAACAATAACAACAACCTCGGGATCGA 2681
Db 2641 AACAAACAACAATTAACAATAACAACAACCTCGGGATCGA 2681

RESULT 14
AX378208
LOCUS AX378208 6806 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 194 from Patent WO0206306.
ACCESSION AX378208
VERSION AX378208.1 GI:19574061
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Yan, R., Tomasselli, A.G., Gurney, M.E., Emmons, T.L., Bienkowski, M.J.
and Heinrikson, R.L.
TITLE Substrates and assays for g(b)-secretase activity

JOURNAL Patent: WO 0206306-A 194 24-JAN-2002;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1..6806
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="fusion protein comprising a maltose binding protein
with 125 amino"

ORIGIN
Query Match 79.9%; Score 2636.8; DB 6; Length 6806;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2638; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCGACACATCAATGATGTCAAAACCTTTCCGCGTATGGCATGATAGCCCGGAAGAGA 60
Db 1 CCGACACATCAATGATGTCAAAACCTTTCCGCGTATGGCATGATAGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGCTGGTGAATGTGAAACAGTAACTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGCTGGTGAATGTGAAACAGTAACTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGCTTTCCGCGTGTGAACAGGCGAGCCAGCTTTCTGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGCTTTCCGCGTGTGAACAGGCGAGCCAGCTTTCTGCGAAA 180
QY 181 CGCGGAAAAGTGAAGCGCGGATGCGGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAGTGAAGCGCGGATGCGGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGCTGCTGATTTCCGCGTGGCGTGGCACTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGCTGCTGATTTCCGCGTGGCGTGGCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATATCTCGCGCGGATCACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATATCTCGCGCGGATCACTGGGTGCCAGCG 360
QY 361 TGTGTGTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGCGCGTGCACAATC 420
Db 361 TGTGTGTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGCGCGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATCTCCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATCTCCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGCGCTTATTTCTTGATGTCTTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGCGCTTATTTCTTGATGTCTTGACCAGA 540
QY 541 CACCCATCAACAGTATTTCTCCCATGAGACGGTACGCGACTGGGGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTCTCCCATGAGACGGTACGCGACTGGGGGTGGAGCATC 600
QY 601 TGTGTCATTTGGGTACACGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCGTCTCGG 660
Db 601 TGTGTCATTTGGGTACACGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCGTCTCGG 660
QY 661 CGGCTCTGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGGACTGAGTGGCTGCTGCGTTTCCAAACAAACCATGCAAAATCTGA 780
Db 721 CGGAACGGGAAGCGGACTGAGTGGCTGCTGCGTTTCCAAACAAACCATGCAAAATCTGA 780
QY 781 ATGAGGCAATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGGCGCTGGGCGCAA 840
Db 781 ATGAGGCAATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGGCGCTGGGCGCAA 840
QY 841 TGGCGCCCAATTCAGGATCGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 900
Db 841 TGGCGCCCAATTCAGGATCGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 900

Db 841 TGGCGCCATTACGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACACGAGCGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAACACGAGCGGAGCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGGACAGAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGGACAGAGGTTT 1140
Qy 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCAACCAATGCTTCTGGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGATTAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGGTT 1380
Qy 1381 CTGGCAATATTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAATATTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTTAGGTGTTTTCAGGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTTAGGTGTTTTCAGGA 1500
Qy 1501 GCACCTTACCAACAGGACCATAGATTATGAAGCTGAAGAGTTAAACTGCTATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTATGAAGCTGAAGAGTTAAACTGCTATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAATCGGCGATGGCCCTGACATTAATCTTCTGGGCACAGCCGCTTGGTGCGTAC 1740
Db 1681 GCGCAATCGGCGATGGCCCTGACATTAATCTTCTGGGCACAGCCGCTTGGTGCGTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGCAAGCGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGCAAGCGTTCCAGGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCCGTACGTTTACACGCGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCCGTACGTTTACACGCGCAAGCTGATGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGGCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGGCGAAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTCGATTAAGCACTGAAGCGGAAGGTTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATTAAGCACTGAAGCGGAAGGTTAAGAGCGCGCTGATGTTCAAC 1980

Qy 1981 CTGCAAGAACCGTACTTACCTTGGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTTGGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGCAAGTAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGCAAGTAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGAGAAGCTGCTCTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTTG 2220
Db 2161 TCCATCGAGAAGCTGCTCTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTTG 2220
Qy 2221 GCATGGTCCAAATCGACACGACGCAAAAGTGAATTAATGGTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACGACGCAAAAGTGAATTAATGGTGTAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGTCACACCATCAAAACCGTTTGGTGGCTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTTGGTGGCTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCTCGAAAACCTATCTGCTGACTGATGAAGGTCGT 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCTCGAAAACCTATCTGCTGACTGATGAAGGTCGT 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGTGGTGCCGTAGCGCTGAGCTGAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGTGCCGTAGCGCTGAGTCTTACGAGGAAG 2460
Qy 2461 TTGCGAAAGATCCACGCTATTGCGCCCAACCATGGAACAGCCCGAGAAAGTGAATCATG 2520
Db 2461 TTGCGAAAGATCCACGCTATTGCGCCCAACCATGGAACAGCCCGAGAAAGTGAATCATG 2520
Qy 2521 CCGAACAATCCCGAGATGTCGGTTTTTGGTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACAATCCCGAGATGTCGGTTTTTGGTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTCAGACTGTCGATGACGCTGGAAGAGCGGAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTCAGACTGTCGATGACGCTGGAAGAGCGGAGACTTAATTCGAGCTCG 2640

RESULT 15
AF031088
LOCUS AF031088 6706 bp DNA circular SYN 07-MAY-2001
DEFINITION Shuttle vector pMAL-pIII, complete sequence.
ACCESSION AF031088
VERSION AF031088.1 GI:2623821
KEYWORDS Shuttle vector pMAL-pIII
SOURCE Shuttle vector pMAL-pIII
ORGANISM Shuttle vector pMAL-pIII
other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6706)
AUTHORS Zwick,M.B., Bonnycastle,L.L., Noren,K.A., Venturini,S., Leong,E.,
Barbas,C.F. III, Noren,C.J. and Scott,J.K.
TITLE The maltose-binding protein as a scaffold for monovalent display of
peptides derived from phage libraries
JOURNAL Anal. Biochem. 264 (1), 87-97 (1998)
MEDLINE 99002881
PUBMED 9784192
REFERENCE 2 (bases 1 to 6706)
AUTHORS Noren,K.A. and Noren,C.J.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road,
Beverly, MA 01915, USA
FEATURES
source
1..6706
/organism="Shuttle vector pMAL-pIII"
/mol_type="genomic DNA"

```
/db_xref="taxon:69327"
/note="derived from commercially available vector pMal-p2
(New England Biolabs)"
15..1163
/gene="lacI"
/note="E. coli gene"
15..46
/gene="lacI"
81..1163
/gene="lacI"
/note="transcriptional repressor"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAB86558.1"
/db_xref="GI:2623822"
/translation="MKPVTLYDAEYAGVSYQTVSRVUNQASHVSAKTRKVEAMAB
LNYIPNVAQQLAKQSLIGVATSSALHAPSQIVAAIKSRADOLGASVSVMSVRS
GVKCAAHVNLQKRVSLGIINYPDDQDAIAVEAACTNYPALFLVSDQTPINSII
FSHEDGTRLGVHILVAGHQIALLAGLSVSARLRLAGWHKYLTRNQIPIAEREG
DWSMSGFQQTOMLNEGIVPTAMLVANDQMALGMRAITESGLRVGADI.SVVGYYDT
EDSCYIPPLTTIKQDFRLTGQTSVDRLLQSQCQAVKNGQLLPVSLVKRKTTLAPNT
QTASPRALADSLMLQARQVSRLESQ"
1406..1433
/note="Ptac promoter"
1528..2976
/gene="malE-lacZa fusion"
/note="E. coli gene fusion"
1528..2976
/gene="malE-lacZa fusion"
/note="MBP-lacZa; MBP modified to include phage M13
secretion signal sequence"
/codon_start=1
/transl_table=11
/product="maltose binding protein-lacZ alpha peptide
fusion protein precursor"
/protein_id="AAB86559.1"
/db_xref="GI:2623823"
/translation="MKKLLFAIPLVVPFYSHSADMKEEGLNVIWNGDKYGLABV
GKPEKDTGTVVEHDKLEKPPQVAATCGDPDII.FWAHDRFGVQAQGLLAEITP
DKAFQDKLPTVTDVAVNGKLIJAYPIAVALSLIYNKDLLENPPKTWESI.PALDEL
KAKGKSALMENLOQPYFTWPLJADGGYAKYKNGYDIDKVGVDNAGAKAGLTFVLD
LIIKNHNAITDYSIAAANKGETANTINGPNAWSNIDTSKNYGVTVLPTFGQPS
KPFVVLASGAINASPNKELEFENLITDEGLEAVNKDKPLGVALKSYBEEELAK
DPR.LAATMENAGKEINPTPQMSAFWYAVRTAVINAASGRQTVDEALDKDAQTNSSN
NNNNNNNNLIEGRISEFGSRVDLQASIALAVLQRRDWNPNVGTQLNRLAAHPFF
ASMRNSEARTDRPSQQLRSLNGEWQLGCFGG"
1528..1581
/gene="malE-lacZa fusion"
/note="M13 secretion signal sequence"
1582..2973
/gene="malE-lacZa fusion"
/product="maltose binding protein-lacZ alpha peptide
fusion protein"
3478..4338
/gene="bla"
/note="R-plasmid ampicillin resistance gene"
3478..4338
/gene="bla"
/note="penicillinase"
/codon_start=1
/transl_table=11
/product="RTEM-1 beta-lactamase"
/protein_id="AAB86560.1"
/db_xref="GI:2623824"
/translation="MSIQHPRVALIPFFAACPVPFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILDSFRPEERSPMSTKVLICGLSVLRDVGAGQQLGRHHYNSNDLVE
YSPVTEKHLTDGTVRELCSAAITMSDNTAANLLITIGGPKSLTFLPHNMGDHVTPL
DRWPELNEAIPNDERDITMPVAMATLRKLLTGELLTLASRQQLIDMWEADKVGKPL
RLADLPAGWFIADKSGAGERGRIIAALPGDKPSRIVVIYTTGSAKMDERNRQIA
EIGASLIKHW"
4381..4893
/standard_name="M13 origin (+ -)"
```

```
ORIGIN
Query Match      79.7%; Score 2631; DB 12; Length 6706;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 60; Gaps 1;

QY      1  CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
DB      1  CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60

QY     61  GTCAATTTCAGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATCCG 120
DB     61  GTCAATTTCAGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATCCG 120

QY    121  GTGCTCTTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCCAGCGTTTCTGCGAAAA 180
DB    121  GTGCTCTTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCCAGCGTTTCTGCGAAAA 180

QY    181  CCGGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTACATTCCCAACCGGTTGGCAC 240
DB    181  CCGGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTACATTCCCAACCGGTTGGCAC 240

QY    241  AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTTGCCCTGC 300
DB    241  AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTTGCCCTGC 300

QY    301  ACGCGCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGTGGCGAGCG 360
DB    301  ACGCGCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGTGGCGAGCG 360

QY    361  TGGTGGTGTGCGATGCTAGAACGAAGCGCGCTGAAAGCTCTAAAGCGGCGGTGCACAATC 420
DB    361  TGGTGGTGTGCGATGCTAGAACGAAGCGCGCTGAAAGCTCTAAAGCGGCGGTGCACAATC 420

QY    421  TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATATCGCTGGATGACAGAGTGCCA 480
DB    421  TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATATCGCTGGATGACAGAGTGCCA 480

QY    481  TTCTGTGGAAAGCTGCTGCACATAATGTTCCGGGTTATTTCTTGATGTTCTTGACCAGA 540
DB    481  TTCTGTGGAAAGCTGCTGCACATAATGTTCCGGGTTATTTCTTGATGTTCTTGACCAGA 540

QY    541  CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB    541  CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

QY    601  TGTCTGCATTTGGGTACCAAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCG 660
DB    601  TGTCTGCATTTGGGTACCAAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCG 660

QY    661  CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB    661  CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY    721  CGGAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB    721  CGGAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780

QY    781  ATGAGGGGATCGTTCCCACTCGCATGCTGTTGCCAAGCATAGATGGCGCTGGGCGCAA 840
DB    781  ATGAGGGGATCGTTCCCACTCGCATGCTGTTGCCAAGCATAGATGGCGCTGGGCGCAA 840

QY    841  TGGCGGCCATTACCGAGTCCGGCTCGCGTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB    841  TGGCGGCCATTACCGAGTCCGGCTCGCGTGGTGGCGATATCTCGGTAGTGGGATACG 900

QY    901  AGGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960
DB    901  AGGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960

QY    961  GCTCTGCTGGGGCAAAACAGCGTGGACCGCTTGGTGGCACTCTCTCAGGGCCAGGCGTGA 1020
DB    961  GCTCTGCTGGGGCAAAACAGCGTGGACCGCTTGGTGGCACTCTCTCAGGGCCAGGCGTGA 1020
```

```
QY 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAAACACACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAAACACACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCAGAAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTCGACCAATGCTTCTGGCG 1260
Db 1201 GCAGAAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTCGACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGAGCATCGGAAGCTGTGGTATGCTGTGCAGGTCGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGCGAGCATCGGAAGCTGTGGTATGCTGTGCAGGTCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACATCCCGTCTCGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACATCCCGTCTCGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAATTAATCATCGGTCGTAATATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAATTAATCATCGGTCGTAATATGTGTGA 1440
QY 1441 ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCGCTGTTAGTGTTTTTCAG 1500
Db 1441 ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCGCTGTTAGTGTTTTTCAG 1500
QY 1501 GCACCTTCAACCAAGGACCATAG----- 1524
Db 1501 GCACCTTCAACCAAGGACCATAGCATATGAAAAATTAATTATTCGCAATTCCTTTAGTG 1560
QY 1525 -----ATTATGAAAACTGAAGAGGTAATACTGGTAATCTGG 1560
Db 1561 GTACCTTTCTAATCTCACTCGGCCGATATGAAAAATCGAAGAGGTAATACTGGTAATCTGG 1620
QY 1561 ATTTACGCGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGNAATTCGAGAAAGAT 1620
Db 1621 ATTTACGCGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGNAATTCGAGAAAGAT 1680
QY 1621 ACCGNAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGGTT 1680
Db 1681 ACCGNAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGGTT 1740
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTTGGTGCTAC 1800
Db 1741 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTTGGTGCTAC 1800
QY 1741 GCTCAATCTGGCCTGTTGGCTGAATATCAACCGGCAAGCGTTCCAGGCAAGCTGTAT 1800
Db 1801 GCTCAATCTGGCCTGTTGGCTGAATATCAACCGGCAAGCGTTCCAGGCAAGCTGTAT 1860
QY 1801 CCGTTTACCTGGGATGGCGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1861 CCGTTTACCTGGGATGGCGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1920
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAGATCTGCTCGGAAACCCGCAAAAACCTGGGAA 1920
Db 1921 GAAGCGTTATCGCTGATTTTAAACAAGATCTGCTCGGAAACCCGCAAAAACCTGGGAA 1980
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAGGTAAGAGCGCTGATGTTCAAC 1980
Db 1981 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAGGTAAGAGCGCTGATGTTCAAC 2040
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGACGGGGTTATCGGTTCAAG 2040
Db 2041 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGACGGGGTTATCGGTTCAAG 2100
```

```
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGTGGATACGCTGGCGCGAAGCG 2100
Db 2101 TATGAAACCGCAAGTACGACATTAAGACGTGGCGTGGATACGCTGGCGCGAAGCG 2160
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2161 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2220
QY 2161 TCCATTCGAGAAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAAGCGGCCGTGG 2220
Db 2221 TCCATTCGAGAAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAAGCGGCCGTGG 2280
QY 2221 GCATGGTCCACATCGACACAGCAAAAGTGAATTAATGGTGTAACGGTACTCCGACCTTC 2280
Db 2281 GCATGGTCCACATCGACACAGCAAAAGTGAATTAATGGTGTAACGGTACTCCGACCTTC 2340
QY 2281 AAGGCTCAACCATCCAAACCGTTTCTGGCGTGTGAGCGAGGATTAATAAGCGGCCAGT 2340
Db 2341 AAGGCTCAACCATCCAAACCGTTTCTGGCGTGTGAGCGAGGATTAATAAGCGGCCAGT 2400
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGTGACTGATGAAGGTCG 2400
Db 2401 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGTGACTGATGAAGGTCG 2460
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCGTAGCGCTGNAAGTCTTACGAGGAAG 2460
Db 2461 GAAGCGGTTAATAAGACAAACCGCTGGTGGCGTAGCGCTGNAAGTCTTACGAGGAAG 2520
QY 2461 TTGGCGAAAGATCCACGCTATTTCGCGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG 2520
Db 2521 TTGGCGAAAGATCCACGCTATTTCGCGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG 2580
QY 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGATGATCAACGCC 2580
Db 2581 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGATGATCAACGCC 2640
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGACACTAATTCGAGCTCG 2640
Db 2641 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGACACTAATTCGAGCTCG 2700
QY 2641 AACACAAACAAATAACAAATAACAAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
Db 2701 AACACAAACAAATAACAAATAACAAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2760
QY 2701 GGATCCTCT 2709
Db 2761 GGATCCTCT 2769
```

Search completed: August 1, 2005, 17:21:57
Job time : 9284 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-15
Perfect score: 3300
Sequence: 1 ccgacacatgaatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gsel:*
- 9: gb_gsel:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	778	23.6	794	5	BQ751936	BQ751936 EST632499
2	770.6	23.4	779	7	CK394523	CK394523 hggada4D0
3	768	23.3	769	7	CK118014	CK118014 218n06.p1
4	762	23.1	793	5	BQ751655	BQ751655 EST632218
5	755.4	22.9	806	5	BQ751220	BQ751220 EST631783
6	684.8	20.8	689	6	CB863341	CB863341 HH04A08y
7	676.4	20.5	712	1	AL038548	AL038548 DKF2p5566B
8	676	20.5	710	8	AY080106	AY080106 AY080106
9	647.4	19.6	660	1	AL044483	AL044483 DKF2p434I
10	641	19.4	641	4	BJ063872	BJ063872 BJ063872
11	615.8	18.7	624	6	CB863314	CB863314 HH07B20y
12	612	18.5	794	1	AL045353	AL045353 DKF2p434B
13	603	18.3	752	9	CL655575	CL655575 PRI0124a-
14	589.8	17.9	608	8	AY080094	AY080094 AY080094
15	586.4	17.8	589	1	AL037742	AL037742 DKF2p564C
16	581.2	17.6	601	1	AV594538	AV594538 AV594538
17	579	17.5	579	1	AL043868	AL043868 DKF2p434C
18	570	17.3	782	7	CK394246	CK394246 hggada1B1
19	569	17.2	774	1	AL045337	AL045337 DKF2p434A
20	563.2	17.1	585	9	CL658017	CL658017 PRI0130b
21	562.6	17.0	723	1	AL039416	AL039416 DKF2p434L
22	549.4	16.6	753	1	AL045341	AL045341 DKF2p434B
23	546.8	16.6	782	1	AL037051	AL037051 DKF2p564I
24	545.4	16.5	547	6	CA895956	CA895956 B0195F10-

C 25	545	16.5	648	7	CK394349	CK394349 hggada2D0
C 26	540.2	16.4	569	1	AL039077	AL039077 DKF2p5566G
C 27	534.4	16.2	644	6	CB862524	CB862524 HH04B08u
C 28	534	16.2	734	1	AL048427	AL048427 DKF2p556H
C 29	529	16.0	740	1	AL042909	AL042909 DKF2p434J
C 30	528	16.0	720	6	CA881974	CA881974 K0994C10-
C 31	523.2	15.9	528	1	AL038811	AL038811 DKF2p5566O
C 32	519.4	15.7	521	6	CA890032	CA890032 B0156F05-
C 33	519	15.6	607	6	CB862099	CB862099 HH08A02y
C 34	516.4	15.6	518	6	CA886082	CA886082 B0124F04-
C 35	510.2	15.5	686	1	AL044407	AL044407 DKF2p434D
C 36	509.4	15.4	511	6	CA895436	CA895436 B0192A07-
C 37	498	15.1	498	1	AL039076	AL039076 DKF2p5566G
C 38	494.8	15.0	1025	1	AL038025	AL038025 DKF2p5566C
C 39	487.6	14.8	498	7	CK394397	CK394397 hggada2H1
C 40	477.6	14.5	491	6	CB862715	CB862715 HH03F14u
C 41	465.8	14.1	679	1	AL039128	AL039128 DKF2p5566K
C 42	462.8	14.0	490	1	AL039589	AL039589 DKF2p434D
C 43	458	13.9	458	6	CA887583	CA887583 B0137H04-
C 44	456.4	13.8	471	1	AL039649	AL039649 DKF2p434G
C 45	449.2	13.6	612	6	CB863409	CB863409 HH04K05w

ALIGNMENTS

RESULT 1
BQ751936
LOCUS BQ751936 DSCT Colletotrichum trifolii cDNA clone pDSCT8-67, mRNA 794 bp mRNA linear EST 18-JUL-2002
DEFINITION EST632499 DSCT Colletotrichum trifolii cDNA clone pDSCT8-67, mRNA
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 794)
AUTHORS Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
Cheung,F. and Fraser,C.M.
TITLE ESTs from mycelia of Colletotrichum trifolii race 1
JOURNAL Unpublished (2002)
COMMENT Other ESTs: EST632498
Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debby@puccini.crl.umn.edu
TIGR sequence name: MTSAH67TV More information is available at:
www.medicago.org
Seq primer: (gtA AtA CgA Ctc Act AtA ggg C).
Location/Qualifiers
1. 794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSCT8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2%glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"
/note="vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2ep2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce

FEATURES

source

E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

```

Query Match      23.6%; Score 778; DB 5; Length 794;
Best Local Similarity 98.7%; Pred. No. 4.9e-219;
Matches 784; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 3 GACACCATCGAATGGTCAAAACCTTCGCGGTATGCGATGATAGCCCGGAAGAGAGT 62
Db 1 GACACTTTCGAATGGCGCAAAACCTTCGCGGTATGCGATGATAGCCCGGAAGAGAGT 60
Qy 63 CAATTCAGGGTGGTAATGTGAACACAGTAACTATACGATGTCGAGATATGCCGGT 122
Db 61 CAATTCAGGGTGGTAATGTGAACACAGTAACTATACGATGTCGAGAGATATGCCGGT 120
Qy 123 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACACAGCCAGCCACGTTTCTGCCAAAACG 182
Db 121 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACACAGCCAGCCACGTTTCTGCCAAAACG 180
Qy 183 CGGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTAATTCATTCACACCGGTGGCAAA 242
Db 181 CGGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTAATTCATTCACACCGGTGGCAAA 240
Qy 243 CAACCTGGCGGCAACAGTGTCTCATTTGGCGTTCGCACTTCAGTTCGGCCCTGCAC 302
Db 241 CAACCTGGCGGCAACAGTGTCTCATTTGGCGTTCGCACTTCAGTTCGGCCCTGCAC 300
Qy 303 GCGCCGTGCGAAATGTGCGGCGGATTAATCTCGCGCGATCAATCTGGGTGCCAGCGTG 362
Db 301 GCGCCGTGCGAAATGTGCGGCGGATTAATCTCGCGCGATCAATCTGGGTGCCAGCGTG 360
Qy 363 GTGGTGTGATGTAGAACAGCGGCTGGAAGCCCTGTAAGCGGCGGTGCACAACTCTT 422
Db 361 GTGGTGTGATGTAGAACAGCGGCTGGAAGCCCTGTAAGCGGCGGTGCACAACTCTT 420
Qy 423 CTGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCGAGTGCATTT 482
Db 421 CTGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCGAGTGCATTT 480
Qy 483 GCTGTGGAAGCTGCTGCATTAATGTTTCCGCGGTATTTCTTGATGTTCTTGACCCAGACA 542
Db 481 GCTGTGGAAGCTGCTGCATTAATGTTTCCGCGGTATTTCTTGATGTTCTTGACCCAGACA 540
Qy 543 CCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGGCTGGAGCATCTG 602
Db 541 CCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGGCTGGAGCATCTG 600
Qy 603 GTGCGATTGGGTCAACAGCAAACTCGCGCTGTAGCGGGCCCATTTAAGTTCTGTCTGGCG 662
Db 601 GTGCGATTGGGTCAACAGCAAACTCGCGCTGTAGCGGGCCCATTTAAGTTCTGTCTGGCG 660
Qy 663 CGTCTGGCTGTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 661 CGTCTGGCTGTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720
Qy 723 GAACGGGAAGCGACTGAGTGCATGTCGGGTTTTCACAAACCATGCAATGCTGAAT 782
Db 721 GAACGGGAAGCGACTGAGTGCATGTCGGGTTTTCACAAACCATGCAATGCTGAAT 780
Qy 783 GAGGGCATCGTTCC 796
Db 781 GAGGGCATCGTTCC 794

```

RESULT 2
CK394523/c
LOCUS
DEFINITION hggada4D08 Gland Cell Amplified cDNA Library Heterodera glycines
779 bp mRNA linear EST 30-DEC-2003

cDNA, mRNA sequence.

```

CK394523
CK394523.1 GI:40389794
EST.
Heterodera glycines
Heterodera glycines
Bukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
1 (bases 1 to 779)
Wang,X., Allen,R., Gao,B., Goellner,M., Maier,T., Baum,T.,
Hussey,R. and Davis,E.
Submission: Wang, Allen, Gao, Goellner, Maier, Baum, Hussey, Davis
Unpublished (2003)
Contact: Tom Maier
Department of Plant Pathology, Baum Lab
Iowa State University
351 Bessey Hall, Ames, IA 50011, USA
Tel: 515-294-8854
Fax: 515-294-9420
Email: trmaier@iastate.edu
Heterodera glycines Gland Cell Amplified cDNA Library, single pass
sequence.
Location/Qualifiers
1..779
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell Amplified cDNA Library"
/notes="Organ: gland cell; Vector: pSportII"

```

FEATURES

source

```

1..779
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell Amplified cDNA Library"
/notes="Organ: gland cell; Vector: pSportII"

```

ORIGIN

```

Query Match      23.4%; Score 770.6; DB 7; Length 779;
Best Local Similarity 99.4%; Pred. No. 7.6e-217;
Matches 773; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 87 CCAGTAACGTTATACGATGTCGAGAGTATGCCGCTGTCTCTTATCAGACCGTTTCCCGC 146
Db 779 CCAGTAACGTTATACGATGTCGAGAGTATGCCGCTGTCTCTTATCAGACCGTTTCCCGC 720
Qy 147 GTGGTGAACAGCCAGCCAGCCACGTTTCTGCGAAACCGGGAAAAAGTGGAAAGCGGAGT 206
Db 719 GTGGTGAACAGCCAGCCAGCCAGCTTTCTGCGAAACCGGGAAAAAGTGGAAAGCGGAGT 660
Qy 207 GCGGAGCTGAATTACATTCACACCGCTGGGCAACAACCTGCGGGCAACAAGTCTGTTG 266
Db 659 GCGGAGCTGAATTACATTCACACCGCTGGGCAACAACAACCTGCGGGCAACAAGTCTGTTG 600
Qy 267 CTGATTGGGCTTGCCACCTCCAGTCTGGCCCTGACCGCGCTCGCAAAATTTGTCGGGCG 326
Db 599 CTGATTGGGCTTGCCACCTCCAGTCTGGCCCTGACCGCGCTCGCAAAATTTGTCGGGCG 540
Qy 327 ATTAAATCTCGGCCCATCAACTGGGTGCGAGGTGGTGGTGTGATGTTAGAACGAAGC 386
Db 539 ATTAAATCTCGGCCCATCAACTGGGTGCGAGGTGGTGGTGTGATGTTAGAACGAAGC 480
Qy 387 GGGCTCGAAGCCCTGTAAGCGCGGTGCAATCTTCTCGCGCAACCGCTCAGTGGGCTG 446
Db 479 GGGCTCGAAGCCCTGTAAGCGCGGTGCAATCTTCTCGCGCAACCGCTCAGTGGGCTG 420
Qy 447 ATCATTAACTATCCGCTGGATGACCAAGATGCCATTGCTGTGGAACTGCTCTCACTAAT 506
Db 419 ATCATTAACTATCCGCTGGATGACCAAGATGCCATTGCTGTGGAACTGCTCTCACTAAT 360
Qy 507 GTTCCGGCGTTATTTCTTGATGTCTGACCGAGACCCATCAACAGTATATTTCTTCC 566
Db 359 GTTCCGGCGTTATTTCTTGATGTCTGACCGAGACCCATCAACAGTATATTTCTTCC 300
Qy 567 CATGAAGACGGTACGGACTGGGCGTGGAGCATCTGTCGCAATTTGGGTCAACAGCAATC 626
Db 299 CATGAAGACGGTACGGACTGGGCGTGGAGCATCTGTCGCAATTTGGGTCAACAGCAATC 240

```

Qy 627 GCGCTGTTAGCGGCCCAATTAACTTCTGTCTCGCGCGGTCTGGCTGGCTGGCAT 686
Db 239 GCGCTGTTAGCGGCCCAATTAACTTCTGTCTCGCGCGGTCTGGCTGGCTGGCAT 180
Qy 687 AAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGAGCTGGAGTGCC 746
Db 179 AAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGAGCTGGAGTGCC 120
Qy 747 ATGTCGGGTTTTCAACAAACCAATGCAATGCTGAATGAGGCATCGTTPCCCACTGCGGATG 806
Db 119 ATGTCGGGTTTTCAACAAACCAATGCAATGCTGAATGAGGCATCGTTPCCCACTGCGGATG 60
Qy 807 CTGGTTGCAACGATCAGATGGCGCTGGCGGCAATGCGCGCAATACCGAGTCCGGGC 864
Db 59 CTGGTTGCAACGATCAGATGGCGCTGGCGGCAATGCGCGCAATACCGAGTCCGGGC 2

RESULT 3
CK118014 769 bp mRNA linear EST 01-JUN-2004
LOCUS 218n06.pl AtM1 Arabidopsis thaliana cDNA clone MPMPGP2011N06218
DEFINITION 5-PRIME, mRNA sequence.

ACCESSION CK118014
VERSION CK118014.1 GI:47828330
KEYWORDS EST.
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 769)
Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.
AUTHORS Generation of a cDNA expression library from Arabidopsis
TITLE inflorescence meristem
JOURNAL Unpublished (2003)
COMMENT Contact: Birgit Kersten
Plant Protein Chip Group, Department Lehrach
Max-Planck-Institute for Molecular Genetics
Inhestr. 73, D-14195 Berlin, Germany
Tel.: +49 (0) 30/84131648
Fax: +49 (0) 30/84131128
Email: Kersten@molgen.mpg.de
Insert Length: 769 Std Error: 0.00
Plate: 218 row: N column: 6
Seq primer: PQE65.

FEATURES
source Location/Qualifiers
1..769
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:956015"
/db_xref="taxon:3702"
/clone="MPMPGP2011N06218"
/tissue_type="inflorescence meristem"
/dev_stages="about one week after bolting"
/lab_host="E. coli SCS-1/pSE111"
/clone_lib="AtM1"
/note="Vector: pQR-30NAST-attB (AY386205); Site 1: SalI;
Site 2: NotI; About 1 week after bolting, cDNA synthesis
using SuperscriptTM-system (Invitrogen) with an
oligo(dT)-primer containing NotI restriction site and a
Sali adapter. The main library (plate numbers begin with
1) of 38,000 clones was rearrayed into the sublibrary
(plate numbers begin with 201) containing 5,000 putative
expression clones. Average insert size is 1 kb. Note: The
rearrayed sublibrary (plate numbers begin with 201) was
sequenced. Library generation and sequencing was granted
in context of GABI-LAPP; data are also accessible at
<https://gabi.rzpd.de>"

ORIGIN

Query Match 23.3%; Score 768; DB 7; Length 769;
Best Local Similarity 100.0%; Pred. No. 4.5e-216;

Matches 768; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 385 GCGCGCTGCGAAGCGGTGTAAGCGCGGTGCAAACTTCTTCGCGCAACGCGTCAAGTGGC 444
Db 1 GCGCGCTGCGAAGCGGTGTAAGCGCGGTGCAAACTTCTTCGCGCAACGCGTCAAGTGGC 60
Qy 445 TGATCATTAACATCCCGTGGATGACACGAGTGCATTCGTGTGGAAGCTGCCTGCACCTA 504
Db 61 TGATCATTAACATCCCGTGGATGACACGAGTGCATTCGTGTGGAAGCTGCCTGCACCTA 120
Qy 505 ATGTTCCGGCGTTATTTCTTGATGCTCTGACACAGACACCCATCAACAGATATATTTTCT 564
Db 121 ATGTTCCGGCGTTATTTCTTGATGCTCTGACACAGACACCCATCAACAGATATATTTTCT 180
Qy 565 CCCATGAAGACGGTACCGGACTGGCGGTGGAGCATCTGGTGCATTTGGGTCAACAGAAA 624
Db 181 CCCATGAAGACGGTACCGGACTGGCGGTGGAGCATCTGGTGCATTTGGGTCAACAGAAA 240
Qy 625 TCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGGCGGTCTGGTCTGGCTGGCTGGC 684
Db 241 TCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGGCGGTCTGGTCTGGCTGGCTGGC 300
Qy 685 ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGCACTGGAGTG 744
Db 301 ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGCACTGGAGTG 360
Qy 745 CCATGTCGGTTTTCAACAAACCAATGCAAAATGCTGAATGAGGCGATCGTTCCTACTGGA 804
Db 361 CCATGTCGGTTTTCAACAAACCAATGCAAAATGCTGAATGAGGCGATCGTTCCTACTGGA 420
Qy 805 TCGTGGTTGCCAAGCATCAGATGCGGTGGCGCAATGCGCGCCATTCACGAGTCCGGGC 864
Db 421 TCGTGGTTGCCAAGCATCAGATGCGGTGGCGCAATGCGCGCCATTCACGAGTCCGGGC 480
Qy 865 TCGCGCTGTGGTGGGATATCTCGGTAGTGGATACGACATACGGAAGACAGCTCATGTT 924
Db 481 TCGCGCTGTGGTGGGATATCTCGGTAGTGGATACGACATACGGAAGACAGCTCATGTT 540
Qy 925 ATATCCCGCGTTAAACCAACCATCAAAAGAGATTTTCGCTCTCGGGCAAAACGACGTGG 984
Db 541 ATATCCCGCGTTAAACCAACCATCAAAAGAGATTTTCGCTCTCGGGCAAAACGACGTGG 600
Qy 985 ACCGCTGTGCTCAACTCTCTCAGGCGCAGGCGGTGAAGGGAATCAGCTGTTCGCCGTCT 1044
Db 601 ACCGCTGTGCTCAACTCTCTCAGGCGCAGGCGGTGAAGGGAATCAGCTGTTCGCCGTCT 660
Qy 1045 CACTGGTGAAAGAAACCAACCCCTTGGCGCCCAATACGCAAAACCGCTCTCCCGCGCGT 1104
Db 661 CACTGGTGAAAGAAACCAACCCCTTGGCGCCCAATACGCAAAACCGCTCTCCCGCGCGT 720
Qy 1105 TGGCCGATTTCATTAATGAGCTGGCAGCAGAGTTTCCCGACTGGAAA 1152
Db 721 TGGCCGATTTCATTAATGAGCTGGCAGCAGAGTTTCCCGACTGGAAA 768
RESULT 4
BO751655 793 bp mRNA linear EST 18-JUL-2002
LOCUS EST6732218 DSCT Colletotrichum trifolii cDNA clone pDSC76-59, mRNA
DEFINITION sequence.
ACCESSION BO751655
VERSION BO751655.1 GI:21907060
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 793)
AUTHORS Samad,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
Cheung,F. and Fraser,C.M.
TITLE ESTs from mycelia of Colletotrichum trifolii race 1
JOURNAL Unpublished (2002)

COMMENT

Other_ESTs: EST632217
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debby@puccini.crl.umn.edu
 TIGR sequence name: MTSAF59TV More information is available at:
 www.medicago.org
 Seq primer: (gtA AtA CgA CtC Act AtA 99g C).

FEATURES

source
 1..793
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDSC16-59"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DHSalpha"
 /clone_lib="DSC1"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DHSalpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 23.1%; Score 762; DB 5; Length 793;
 Best Local Similarity 99.1%; Pred. No. 2.8e-214;
 Matches 777; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
 Qy 3 GACACCATCGAATGGTGCMAAACCTTTCGGGTATGCGATGATGATGCGCGAGAGAGT 62
 Db 1 GACATCTTCGATGGCGCAAAACCTTTCGGGTATGCGATGATGCGCGAGAGAGT 60
 Qy 63 CAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATCGCGT 122
 Db 61 CAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATCGCGT 120
 Qy 123 GTCTCTTATCAGACCGTTTCGGGTGGTGAACAGCGCCAGCCACGTTTCGCGAAACG 182
 Db 121 GTCTCTTATCAGACCGTTTCGGGTGGTGAACAGCGCCAGCCACGTTTCGCGAAACG 180
 Qy 183 CGGGAACAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTACACCGCGTGGCACA 242
 Db 181 CGGGAACAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTACACCGCGTGGCACA 240
 Qy 243 CAATCGCGGCGCAACAGTGGTCTGATTTGGGTGGTCCACCTCCAGTCTGGCCCTGCAC 302
 Db 241 CAATCGCGGCGCAACAGTGGTCTGATTTGGGTGGTCCACCTCCAGTCTGGCCCTGCAC 300
 Qy 303 CGCGCGTCGCAAAATGTCGGGGGATTAATCTCGGCCGATCACTGGGTGCCAGCGTG 362
 Db 301 CGCGCGTCGCAAAATGTCGGGGGATTAATCTCGGCCGATCACTGGGTGCCAGCGTG 360
 Qy 363 GTGGTGTGATGTAGAACAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAATCTT 422
 Db 361 GTGGTGTGATGTAGAACAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAATCTT 420
 Qy 423 CTCGCCAGCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGATGACGAGGATGCATT 482
 Db 421 CTCGCCAGCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGATGACGAGGATGCATT 480

Qy 483 GCTGTGAAGCTGCCTGCACCTAATGTTCCGGCGTTATTTCTTGTATGTTCTGTGACCAGACA 542
 Db 481 GCTGTGAAGCTGCCTGCACCTAATGTTCCGGCGTTATTTCTTGTATGTTCTGTGACCAGACA 540
 Qy 543 CCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGGACTTGGCGGTGGAGCATCTG 602
 Db 541 CCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGGACTTGGCGGTGGAGCATCTG 600
 Qy 603 GTGCATTGGGTCAACAGCAAAATCGCGCTTTAGCGGCGCCATTAAGTTCTGTCTCGCG 662
 Db 601 GTGCATTGGGTCAACAGCAAAATCGCGCTTTAGCGGCGCCATTAAGTTCTGTCTCGCG 660
 Qy 663 CGTCTGGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
 Db 661 CGTCTGGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720
 Qy 723 GAACGGGAAGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGAAT 782
 Db 721 GAACGGGAAGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGAAT 778
 Qy 783 GAGG 786
 Db 779 GAGG 782

RESULT 5
 BO751220
 LOCUS BO751220
 DEFINITION BO751220 806 bp mRNA linear EST 18-JUL-2002
 EST631783 DSC1 Colletotrichum trifolii cDNA clone pDSC13-64, mRNA sequence.
 ACCESSION BO751220
 VERSION BO751220.1 GI:21906625
 KEYWORDS EST.
 SOURCE Colletotrichum trifolii
 ORGANISM Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
 mitosporic Phyllachoraceae; Colletotrichum.
 REFERENCE 1 (bases 1 to 806)
 AUTHORS Samac, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
 Cheung, F. and Fraser, C.M.
 TITLES ESTs from mycelia of Colletotrichum trifolii race 1
 JOURNAL Unpublished (2002)
 COMMENT Other ESTs: EST631782
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debby@puccini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org
 Seq primer: (gtA AtA CgA CtC Act AtA 99g C).

FEATURES

source
 1..806
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDSC13-64"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DHSalpha"
 /clone_lib="DSC1"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid

lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN		Query Match		22.9%; Score 755.4; DB 5; Length 806;		Best Local Similarity		97.4%; Pred. No. 2.5e-212;		Matches 790; Conservative 0; Mismatches 16; Indels 5; Gaps 2;	
Qy	3	GACACCATCGAATGTGTCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGAGT	62			1	GACACTTTCGAATGGCGAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGAGT	60			
Db						63	CAATTTCAGGGTGTGAATGTGAAACACAGTAACGTTATACGATGTGCGAGAGTATGCCGGT	122			
Qy						61	CAATTTCAGGGTGTGAATGTGAAACACAGTAACGTTATACGATGTGCGAGAGTATGCCGGT	120			
Db						123	GTCTCTTATCAGACGGTTTCGGGTGTGTGAACACGCGCAGCAGCTTTCTGCGAAACG	182			
Qy						121	GTCTCTTATCAGACGGTTTCGGGTGTGTGAACACGCGCAGCAGCTTTCTGCGAAACG	180			
Db						183	CGGGAATAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGGTGSCACAA	242			
Qy						181	CGGGAATAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGGTGSCACAA	240			
Db						243	CAACTGGCGGCAACACAGTCGTGTGATTTGGCGTGGCCACTCCAGTCGTGCCCTGCAC	302			
Qy						241	CAACTGGCGGCAACACAGTCGTGTGATTTGGCGTGGCCACTCCAGTCGTGCCCTGCAC	300			
Db						303	CGCGCGTCGCAAAATGTTCGGCGCGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCGTG	362			
Qy						301	CGCGCGTCGCAAAATGTTCGGCGCGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCGTG	360			
Db						363	GTGTGTGTCGATGTAGACGAGCGGCTCGAAGCCTGTAAGCGCGGTGCACAACTCTT	422			
Qy						361	GTGTGTGTCGATGTGTAGACGAGCGGCTCGAAGCCTGTAAGCGCGGTGCACAACTCTT	420			
Db						423	CTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACGAGATGCCATT	482			
Qy						421	CTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACGAGATGCCATT	480			
Db						483	GCTGTGGAAGTGTGCTGCACATAATGTTCCGGGTATTTCTTGATGTCTTGACACAGACA	542			
Qy						481	GCTGTGGAAGTGTGCTGCACATAATGTTCCGGGTATTTCTTGATGTCTTGACACAGACA	540			
Db						543	CCCATCAACAGTATTTTCTCCCATGACGCGTACGCGACTGGGCGTGGAGCATCTG	602			
Qy						541	CCCATCAACAGTATTTTCTCCCATGACGCGTACGCGACTGGGCGTGGAGCATCTG	600			
Db						603	GTCCGATTGGGTTCACGACCAATCCGCTGTTAGCGGCGCCCAATTAAGTTCTGTCTCGGCG	662			
Qy						601	GTCCGATTGGGTTCACGACCAATCCGCTGTTAGCGGCGCCCAATTAAGTTCTGTCTCGGCG	660			
Db						663	CGTCTGCTGTGGTGTGCTGCATAAATATCTACTCGCAATCAAAATTCAGCCGATACGC	722			
Qy						661	CGTCTGCTGTGGTGTGCTGCATAAATATCTACTCGCAATCAAAATTCAGCCGATACGC	720			
Db						723	GAAACGGGAAGCGACTGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGAAT	782			
Qy						721	GAAACGGGAAGCGACTGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGAAT	777			
Db						783	GAGGCGATCGTTCCCACTGCCATGCTGGTTG	813			
Qy						778	GAGGCGATCGTTCCCACTGCCATGCTGGTTG	806			

RESULT 6
CB863541/c
LOCUS HH04A08y HH Hordeum vulgare cDNA clone HH04A08 3-PRIME, mRNA
DEFINITION CB863541 689 bp mRNA linear EST 22-APR-2003
sequence.

CB863541
CB863541.1 GI:30058100
EST.
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 689)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
Graner,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 689 Std Error: 0.00
Plate: 4 row: A column: 8
Seq primer: SP6.
Location/Qualifiers
1..689
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555198"
/db_xref="taxon:4513"
/clone="HH04A08"
/tissue_type="coleoptile"
/dev_stage="coleoptile, 1 day old"
/lab_host="DH10B"
/clone_lib="HH"
/note="Vector: pSPORT; Site 1: SalI (5-end of cDNA);
Site 2: NotI (3-end of cDNA); Due to the cloning system
used blue/white selection for recombinants is not 100 %
reliable. Average insert size is 1.3 kb."

ORIGIN		Query Match		20.8%; Score 684.8; DB 6; Length 689;		Best Local Similarity		99.6%; Pred. No. 2.3e-191;		Matches 686; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Qy	21	AAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGAGTCAATTTCAGGTGTGTAAT	80			689	AAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGAGTCAATTTCAGGTGTGTAAT	630			
Db						81	GTGAAACCAAGTAACTTATACGATGTGCGAGATATGCGCGTCTCTTATCAGACCGTT	140			
Qy						81	GTGAAACCAAGTAACTTATACGATGTGCGAGATATGCGCGTCTCTTATCAGACCGTT	140			
Db						629	GTGAAACCAAGTAACTTATACGATGTGCGAGATATGCGCGTCTCTTATCAGACCGTT	570			
Qy						141	TCCCGCGTGTGAACCAAGCGCCAGCCAGTTCTCGGAAACCGCGGAAAGTGGAAAGCG	200			
Db						569	TCCCGCGTGTGAACCAAGCGCCAGCCAGTTCTCGGAAACCGCGGAAAGTGGAAAGCG	510			
Qy						201	CGCATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCAACAACCTGCGCGGCAACAG	260			
Db						509	CGCATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCAACAACCTGCGCGGCAACAG	450			
Qy						261	TCGTTGTGATTTGGCGTTTGCACCTTCCAGTGTGCGCCCTGCAACGCGCGCTCGCAAAATTGTC	320			
Db						449	TCGTTGTGATTTGGCGTTTGCACCTTCCAGTGTGCGCCCTGCAACGCGCGCTCGCAAAATTGTC	390			
Qy						321	CGGCGGATTAATCTCGCGCGATCACTGGGTGCCAGTGTGTTCCGATGTTAGAA	380			
Db						389	CGGCGGATTAATCTCGCGCGATCACTGGGTGCCAGTGTGTTCCGATGTTAGAA	330			
Qy						381	CGAAGCGCGTTCGAAGCCTGTAAAGCGCGGTGACAAATCTTCTCGCGCAACCGCTCAGT	440			
Db						329	CGAAGCGCGTTCGAAGCCTGTAAAGCGCGGTGACAAATCTTCTCGCGCAACCGCTCAGT	270			

```

QY 441 GGGCTGATCATTAACATATCCGCTGGATGACAGAGATGCCATTGCTGTGGAAGCTGCCTGC 500
Db 269 GGGCTGATCATTAACATATCCGCTGGATGACAGAGATGCCATTGCTGTGGAAGCTGCCTGC 210
QY 501 ACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACAGACACCCATCAACAGTATTATT 560
Db 209 ACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACAGACACCCATCAACAGTATTATT 150
QY 561 TTCTCCCATGAAGACGGTACGCACTGGGCGTGGAGCATCTGGTCGCAATGGGTCAACAG 620
Db 149 TTCTCCCATGAAGACGGTACGCACTGGGCGTGGAGCATCTGGTCGCAATGGGTCAACAG 90
QY 621 CAAATCGCGCTGTAGCGGCCCAATTAAGTCTGCTCGGCGCTGTGCGTCTGGCTGGC 680
Db 89 CAAATCGCGCTGTAGCGGCCCAATTAAGTCTGCTCGGCGCTGTGCGTCTGGCTGGC 30
QY 681 TGGCATAAATAATCTCACTCGCAATCAAT 709
Db 29 TGGCATAAATAATCTCACTCGCAATCAAT 1

RESULT 7
AL038548/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL038548 712 bp mRNA linear EST 04-SEP-2003
DKFZp566E0346_r1 566 (synonym: hfkd2) Homo sapiens cDNA clone
DKFZp566E0346_5, mRNA sequence.
AL038548
AL038548.1 GI:5407738
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 712)
Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Ottenwaelder, et al.)
Unpublished (1995)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZp566E0346) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. 712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp566E0346"
/tissue_type="kidney"
/dev_stage="fetal"
/lab_host="X1-2blue"
/clone_lib="566 (synonym: hfkd2)"
/notes="Vector: pAMPI; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match 20.5%; Score 676.4; DB 1; Length 712;
Best Local Similarity 99.6%; Pred. No. 7.2e-189;
Matches 699; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 33 GGTATGTCATGATAGCGCCCGGAAGAGAGTCAATTGAGGTGGTGAATGTGAACCAAGTA 92
Db 712 GGTATGTCATGATAGCGCCCGGAAGAGAGTCAATTGAGGTGGTGGTGAATGTGAACCAAGTA 653
QY 93 ACCTTATACATGTCGACAGATGCGGTGCTCTTATCAGACCGCTTCCGCGTGGTG 152

```

```

Db 652 ACGTTATACGATGTCGACAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGGTGGTG 593
QY 153 AACCAGGCCAGCCACGTTTCTGCGAAACCGGGG-AAAAAGTGAAGCGGCGATGGCGGA 211
Db 592 AACCAGGCCAGCCACGTTTCTGCGAAACCGGGGAAAAAGTGAAGCGGCGATGGCGGA 533
QY 212 GCTGAATTAACATTCGCAACCGCGTGGGCAACAACCTGGCGGGCAACAGTGTGCTGTAT 271
Db 532 GCTGAATTAACATTCGCAACCGCGTGGGCAACAACCTGGCGGGCAACAGTGTGCTGTAT 473
QY 272 TGGCGTTGCGCACTCCAGTCTGCGCCCTGCAC-GGGCGGTGCGCAATTTGTGCGGCGATT 330
Db 472 TGGCGTTGCGCACTCCAGTCTGCGCCCTGCACGGCGCGTGCAGAAATTTGTGCGGCGATT 413
QY 331 AATCTCGCGCGCATCAACTCGGTGCCAGCGTGGTGTGTCGATGCTAGAGAACGAGCGGCG 390
Db 412 AATCTCGCGCGCATCAACTCGGTGCCAGCGTGGTGTGTCGATGCTAGAGAACGAGCGGCG 353
QY 391 TCGAAGCCTCTGAAGCGGCGGTGCACAACTCTTCGCGCAACCGGTCAGTGGGCTGATCA 450
Db 352 TCGAAGCCTCTGAAGCGGCGGTGCACAACTCTTCGCGCAACCGGTCAGTGGGCTGATCA 293
QY 451 TTAACATATCCGCTGGATGACAGAGTGCACATTCGCTGTGGAAGCTGCTGCACATAATGTTTC 510
Db 292 TTAACATATCCGCTGGATGACAGAGTGCACATTCGCTGTGGAAGCTGCTGCACATAATGTTTC 233
QY 511 CGGCGTTATTTCTTGATGCTCTGACAGACACCCATCAACAGTATTATTTCTCCCATG 570
Db 232 CGGCGTTATTTCTTGATGCTCTGACAGACACCCATCAACAGTATTATTTCTCCCATG 173
QY 571 AAGACGGTACGGCACTGGGCGTGGAGCACTGCTGTCGATGGTGGTGCACCAACCAATCGCGC 630
Db 172 AAGACGGTACGGCACTGGGCGTGGAGCACTGCTGTCGATGGTGGTGCACCAACCAATCGCGC 113
QY 631 TGTAGCGGCGCCATTAAAGTTCTGTCGCGCGCTGCTGGCTGCTGGCTGCTGCATAAAT 690
Db 112 TGTAGCGGCGCCATTAAAGTTCTGTCGCGCGCTGCTGGCTGCTGGCTGCTGCATAAAT 53
QY 691 ATCTCACTCGCAATCAAAATTCAGCGATAGCGGAACGGGAAG 732
Db 52 ATCTCACTCGCAATCAAAATTCAGCGATAGCGGAACGGGAAG 11

AV080106 710 bp DNA linear GSS 06-NOV-2002
AY080106 Scripps Pier (La Jolla, CA) uncultured virus community
uncultured marine virus genomic clone SIO51p3G6L, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

AY080106.1 GI:24745302
uncultured marine virus
uncultured marine virus
Viruses; environmental samples.
1 (bases 1 to 710)
Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
Mead,D., Azam,F. and Rohwer,F.
Genomic analysis of uncultured marine viral communities
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
22294988
Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@sunstroke.sdsu.edu
Class: shotgun.
Location/Qualifiers
1. 710
/organism="uncultured marine virus"

```


/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone="SI051p36L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus community"
/note="Marine viruses were isolated from 200 liters of surface seawater using a combination of differential filtration and density-dependent gradient centrifugation. Linker-amplified shotgun libraries were constructed by randomly shearing the total marine viral community DNA, end-repairing, ligating dsDNA linkers to the ends, and amplifying the fragments using Vent DNA polymerase. The resulting fragments were ligated into the pSMART vector and electroporated into MC12 cells (Lucigen; Middleton, WI)."

ORIGIN

Query Match 20.5%; Score 676; DB 8; Length 710;
Best Local Similarity 98.9%; Pred. No. 9.5e-189;
Matches 690; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
Qy 400 GTAAAGCGCGGTGCACAACTCTTCGGCGCAACGGCTCAGTGGGCTGATCAATTAACATATC 459
Db 8 GTGCTGCAGCGGTGCACAACTCTTCGGCGCAACGGCTCAGTGGGCTGATCAATTAACATATC 67
Qy 460 CGCTGGATGACAGATGCCATTTGCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTTAT 519
Db 68 CGCTGGATGACAGATGCCATTTGCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTTAT 127
Qy 520 TTCTTGATGTCCTGACAGACACCCATCAACAGTATTATTTCTCCCATGAAGCGGTA 579
Db 128 TTCTTGATGTCCTGACAGACACCCATCAACAGTATTATTTCTCCCATGAAGCGGTA 187
Qy 580 CGCACTGGGCGTGAGCATCTGTGCGATTTGGGTCAACAGCAAAATCGCGTGTAGCGG 639
Db 188 CGCACTGGGCGTGAGCATCTGTGCGATTTGGGTCAACAGCAAAATCGCGTGTAGCGG 247
Qy 640 GCCATTAAAGTTCTGCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 699
Db 248 GCCATTAAAGTTCTGCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 307
Qy 700 GCAATCAATTCAGCCGATAGCGGAACGGGAGCGACTGGATGCCATGTCGGGTTTC 759
Db 308 GCAATCAATTCAGCCGATAGCGGAACGGGAGCGACTGGATGCCATGTCGGGTTTC 367
Qy 760 AACAAACCATGCAATGCTGAATAGGGCATCGTTCCCACTGCCATGCTGGTTGCCAACG 819
Db 368 AACAAACCATGCAATGCTGAATAGGGCATCGTTCCCACTGCCATGCTGGTTGCCAACG 427
Qy 820 ATCAGATGGCGTGCGGCGCAATGCGCGCATTAACAGATGCGGGCTGCGGTTGGTGGG 879
Db 428 ATCAGATGGCGTGCGGCGCAATGCGCGCATTAACAGATGCGGGCTGCGGTTGGTGGG 487
Qy 880 ATATCTCGGTAGTGGATACGAGATACGAGACAGCTCATGTTATATCCCGCGTTAA 939
Db 488 ATATCTCGGTAGTGGATACGAGATACGAGACAGCTCATGTTATATCCCGCGTTAA 547
Qy 940 CCACCATCAACAGGATTTTCGCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAAC 999
Db 548 CCACCATCAACAGGATTTTCGCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAAC 607
Qy 1000 TCTCTCAGGGCGAGCGGTGAAGGCAATCAAGCTGTTGCCCGTCTCACTGGTGAAGAA 1059
Db 608 TCTCTCAGGGCGAGCGGTGAAGGCAATCAAGCTGTTGCCCGTCTCACTGGTGAAGAA-AA 666
Qy 1060 AAACCACTTGGCGCCCAATAGCGAAACCGCTCTCC 1097
Db 667 AAACCACTTGGCGCCCAATAGCGAAACCGCTCTCC 704

RESULT 9
AL044483/c
LOCUS

DEFINITION DKFZp4341102_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKFZp4341102.3, mRNA sequence.
VERSION AL044483
KEYWORDS AL044483.1 GI:5432701
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Ansong, W., Benes, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Ansong, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wilmann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wilmann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp4341102) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..660
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp4341102"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN
Query Match 19.6%; Score 647.4; DB 1; Length 660;
Best Local Similarity 99.7%; Pred. No. 2.9e-180;
Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 85 AACCAGTAACGTTATACGATGTCGAGAGTATGCGGTGTCTCTTATCAGACCGTTTCC 144
Db 660 AACCAGTAACGTTATACGATGTCGAGAGTATGCGGTGTCTCTTATCAGACCGTTTCC 601
Qy 145 CGGTGGTGAACAGCCAGCCAGCCAGCTTCTTCGGAACACGGGGAAAAAGTGGAGCGGGA 204
Db 600 CGGTGGTGAACAGCCAGCCAGCCAGCTTCTTCGGAACACGGGGAAAAAGTGGAGCGGGA 541
Qy 205 TGGCGGAGCTCAATTTACATTTCCCAACCGCGTGGCAACAACCTGGCGGGCAACAGTCT 264
Db 540 TGGCGGAGCTCAATTTACATTTCCCAACCGCGTGGCAACAACCTGGCGGGCAACAGTCT 481
Qy 265 TGCTGATTGGGTTGCCACCTCCAGTCTGGCCCTGCACGCGCGTCGCAAAATGTCGGGG 324
Db 480 TGCTGATTGGGTTGCCACCTCCAGTCTGGCCCTGCACGCGCGTCGCAAAATGTCGGGG 421
Qy 325 CGATTAAATCTCGCGCCGATCAACTGGGTGTCAGCGTGGTGTGTCGATGGTAGAACGAA 384
Db 420 CGATTAAATCTCGCGCCGATCAACTGGGTGTCAGCGTGGTGTGTCGATGGTAGAACGAA 361
Qy 385 CGCGGCTCGAAGCCCTGTAAGCGGGGTGCACATCTTCTCGCGCAACGCGTCAGTGGGC 444
Db 360 CGCGGCTCGAAGCCCTGTAAGCGGGGTGCACATCTTCTCGCGCAACGCGTCAGTGGGC 301
Qy 445 TGATCATTAATATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTA 504
Db 300 TGATCATTAATATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTA 241
Qy 505 ATGTTCCGGCGTTATTTCTTGATGTCTCTGACAGACACCCATCAACAGTATTTTCTT 564

```

Db      240  ATGTTCCGGGTTATTTCCTTGATCTCTGACACAGACACCATCAACAGTATTATTTCT 181
QY      565  CCATGAAGACGGTACCGACTGGCGGTGAGCATCTGTCGATTTGGGTCAACAGCAA 624
Db      180  CCATGAAGACGGTACCGACTGGCGGTGAGCATCTGTCGATTTGGGTCAACAGCAA 121
QY      625  TCGCGCTGTTAGCGGGGCCCAATTAAGTTCTGTCGCGGGTCTGCGTCTGGCTGGCTGGC 684
Db      120  TCGCGCTGTTAGCGGGGCCCAATTAAGTTCTGTCGCG-GCGTCTGCGTCTGGCTGGCTGGC 62
QY      685  ATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAAGCGGACTGGAGTG 744
Db      61  ATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAAGCGGACTGGAGTG 2
QY      745  C 745
Db      1  C 1

RESULT 10
BJ063872/c
LOCUS
DEFINITION
BJ063872 N1BB Mochii normalized Xenopus tailbud library EST 29-SEP-2003
laevis cDNA clone XL077k02 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 641)
GI:17471074
http://xenopus.nibb.ac.jp.
Location/Qualifiers
1..641
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:9355"
/clone="XL077k02"
/tissue_type="whole embryo"
/dev stage="stage 25"
/clone.lib="N1BB Mochii normalized Xenopus tailbud
library"

FEATURES
source
Query Match 19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred. No. 2.3e-178;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30  CGCGGTATGCATAGCCCGCGAAGAGAGTCAATTCAGGGTGTGATGTGAACCA 89
Db      641  CGCGGTATGCATAGCCCGCGAAGAGAGTCAATTCAGGGTGTGATGTGAACCA 582
QY      90  GTAAAGTTATACAGTGTGCGAGAGTATGCCGGGTCTCTTATCAGACCGTTTCCGCGTG 149
Db      581  GTAAAGTTATACAGTGTGCGAGAGTATGCCGGGTCTCTTATCAGACCGTTTCCGCGTG 522
QY      150  GTGAACCGCCAGCCAGCTTTCTGCGAATAACCGGGGAAAAAGTGGAGCGCGATGGCG 209

```

```

Db      521  GTGAACCGCCAGCCAGCTTTCTGCGAATAACCGGGGAAAAAGTGGAGCGCGATGGCG 462
QY      210  GAGCTGAATTACATTCACAACCGCGTGGGCAACAACCTGGCGGGCAACAGTCTGTTGCTG 269
Db      461  GAGCTGAATTACATTCACAACCGCGTGGGCAACAACCTGGCGGGCAACAGTCTGTTGCTG 402
QY      270  ATTGGCGTTGGCACTCTCCAGTCTGCGCCTGCAACGCGCGGTGGCAAAATTGTCGGCGGCAATT 329
Db      401  ATTGGCGTTGGCACTCTCCAGTCTGCGCCTGCAACGCGCGGTGGCAAAATTGTCGGCGGCAATT 342
QY      330  AATATCTCGCGCGGATCAACTGGGTGCGAGGTGTGTGTCGATGTGTAGAACGAAGCGGC 389
Db      341  AATATCTCGCGCGGATCAACTGGGTGCGAGGTGTGTGTCGATGTGTAGAACGAAGCGGC 282
QY      390  GTCGAAGCCTGTAAGCGCGGTGCAATCTTCTCGCGCAACGCGTCACTGGGCTGATC 449
Db      281  GTCGAAGCCTGTAAGCGCGGTGCAATCTTCTCGCGCAACGCGTCACTGGGCTGATC 222
QY      450  ATTAATCTATCCGTGGATGACAGGATGCCATTTGCTGTGGAAGCTGCTGCACATAATGTT 509
Db      221  ATTAATCTATCCGTGGATGACAGGATGCCATTTGCTGTGGAAGCTGCTGCACATAATGTT 162
QY      510  CCGCGCTTATTTCTTGATGTCTCTGACCAAGACACCATCAACAGTATTATTTTCTCCCAT 569
Db      161  CCGCGCTTATTTCTTGATGTCTCTGACCAAGACACCATCAACAGTATTATTTTCTCCCAT 102
QY      570  GAAGACGGTACGCGACTGGCGGTGGAGCATCTGGTGCATTTGGTCCAGTCCAGCAAAATCGCG 629
Db      101  GAAGACGGTACGCGACTGGCGGTGGAGCATCTGGTGCATTTGGTCCAGTCCAGCAAAATCGCG 42
QY      630  CTGTTAGCGCGCCATTAAGTTCTGTCGCGCGGCTCTCGC 670
Db      41  CTGTTAGCGCGCCATTAAGTTCTGTCGCGCGGCTCTCGC 1

```

RESULT 11
CB863814/c

LOCUS
DEFINITION
CB863814
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CB863814 624 bp mRNA linear EST 22-APR-2003
HH07B20y HH Hordeum vulgare cDNA clone HH07B20 3-PRIME, mRNA
sequence.
CB863814
CB863814.1 GI:30058373
EST.
Hordeum vulgare
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 624)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
Graner,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 7 row: B column: 20
Seq primer: SP6.
Location/Qualifiers
1..624
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555471"
/db_xref="taxon:4513"
/clone="HH07B20"
/tissue_type="coleoptile"

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

FEATURES
source

ORIGIN

ORIGIN	/dev_stage="coleoptile, 1 day old"	
	/lab_host="DH10B"	
TITLE	EST (Duesterhoeft, et al.)	
	Unpublished (1999)	
JOURNAL	Contact: MIPS	
	MIPS	
COMMENT	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany	
	This is the 3' sequence of the clone insert	
FEATURES	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer	
	Research Center (DKFZ); sequenced by Qiagen within the cDNA	
source	sequencing consortium of the German Genome Project. r1 sequence	
	also available.	
1. .794	This clone is available at the RZPD in Berlin.	
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059	
location/Qualifiers	Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.	
	/organism="Homo sapiens"	
/mol_type="mRNA"	/db_xref="taxon:9606"	
	/clone="DKFZp434B075"	
/tissue_type="testis"	/dev_stage="adult"	
	/lab_host="DH10B"	
/clone.lib="434 (synonym: htes3)"	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"	
	ORIGIN	
Query Match	18.5%; Score 612; DB 1; Length 794;	
	Best Local Similarity 99.7%; Pred. NO. 1e-169;	
Matches 634; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	569 TGAGACGGTACGGACCTGGCGGTGGAGCATCTGGTCGCATTTGGTACCGCAATCGC 628	
	794 TGAAGACGGTACGGACCTGGCGGTGGAG-ATCTGGTCGCATTTGGTACCGCAATCGC 736	
QY	629 GCTGTTAGCGGCGCCATTAAAGTTCCTCGCGCGCTCTGGTCTGGCTGGCTGGCATAA 688	
	735 GCTGTTAGCGGCGCCATTAAAGTTCCTCGCGCGCTCTGGTCTGGCTGGCATAA 677	
QY	689 ATATCTCACTCGCAATCAAAATTCAGCCGATACGGAACGGGAAGCGCATGAGATGCCAT 748	
	676 ATATCTCACTCGCAATCAAAATTCAGCCGATACGGAACGGGAAGCGCATGAGATGCCAT 617	
QY	749 GTCCGGTTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTCCCACTGCCATGCT 808	
	616 GTCCGGTTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTCCCACTGCCATGCT 557	
QY	809 GGTTCGCCAACGATCAGATGGCGCTGGGGCGCAATCGCGCCATTACCGAGTCCGGCTGGC 868	
	556 GGTTCGCCAACGATCAGATGGCGCTGGGGCGCAATCGCGCCATTACCGAGTCCGGCTGGC 497	
QY	869 GGTTCGGTCGGATATCTCGGTAGTGGGATACGAGATACCGAAGACAGCTCATGTTATAT 928	
	496 GGTTCGGTCGGATATCTCGGTAGTGGGATACGAGATACCGAAGACAGCTCATGTTATAT 437	
QY	929 CCCGCCGTTAACCCACCATCAAAACAGGATTTTCGCTGCTGGGGCAAAACAGCGGTGGACCG 988	
	436 CCCGCCGTTAACCCACCATCAAAACAGGATTTTCGCTGCTGGGGCAAAACAGCGGTGGACCG 377	
QY	989 CTTGCTGCAACTCTCTCAGGGCGCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACT 1048	
	376 CTTGCTGCAACTCTCTCAGGGCGCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACT 317	
QY	1049 GGTGAAAAAGAAAAACCACTCCGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTTGGC 1108	
	316 GGTGAAAAAGAAAAACCACTCCGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTTGGC 257	
QY	1109 CGATTCAATATGAGCTGGCAAGCAGGTTTCCCGACTGGAAAGCGGCGAGTAGCGCA 1168	
	256 CGATTCAATATGAGCTGGCAAGCAGGTTTCCCGACTGGAAAGCGGCGAGTAGCGCA 197	
QY	1169 ACGCAATTAATGAGTTAGCTCACTCACTATTAGGCAC 1204	
	196 ACGCAATTAATGAGTTAGCTCACTCACTATTAGGCAC 161	
RESULT 12	AL045353	
	DKFZp434B075.e1 434 (synonym: htes3) Homo sapiens cDNA clone	
DEFINITION	DKFZp434B075.3', mRNA sequence.	
	AL045353	
ACCESSION	AL045353.1 GI:54333506	
	EST.	
KEYWORDS	Homo sapiens (human)	
	Homo sapiens	
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 794)	
	Duesterhoeft.A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann,S.	

```

RESULT 13
CL655575/c
LOCUS
DEFINITION
  CL655575 752 bp DNA linear GSS 09-JUL-2004
  PRI0124a_C01 - PRI0124a.B21 (752) Mixed stage fosmid library of P.
  pacificus var. California Pristionchus pacificus genomic, genomic
  survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Pristionchus pacificus
  Pristionchus pacificus
  Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
  Neodiplogasteridae; Pristionchus.
REFERENCE
  1 (bases 1 to 752)
  Srinivasan, J., Otto G.W., Kahlow, U., Geisler, R. and Sommer, R.J.
  AppaDB: an AcetB database for the nematode satellite organism
  Pristionchus pacificus
  Nucleic Acids Res. 32 (1), D421-D422 (2004)
  Contact: Sommer RJ
  Evolutionary Biology
  Max-Planck-Institute for Developmental Biology
  Spemannstr. 37-39, Tuebingen D-72076, Germany
  Tel.: 00497071601371
  Fax: 00497071601498
  Email: raif.sommer@tuebingen.mpg.de
  This library was generated at Caltech, Pasadena, USA and end
  sequenced at Vancouver, Canada.
  Seq primer: T7
  Class: fosmid ends.
  Location/Qualifiers
    1..752
    /organism="Pristionchus pacificus"
    /mol_type="genomic DNA"
    /strain="California"
    /db_xref="taxon:54126"
    /clone_lib="Mixed stage fosmid library of P. pacificus
    var. California"
    /note="Vector: pBpifos-5 Fosmid vector"

ORIGIN
  source
  Query Match 18.3%; Score 603; DB 9; Length 752;
  Best Local Similarity 98.9%; Pred. No. 4.7e-167;
  Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 1996 TTCACCTGGCGCTGATTGCTGTGAC -GGGGGTTATGCGTTCAAGTATGAAACGGCAA 2054
DB 752 TTCACCTGGCGCTGATTGCTGTGACGGGGGTTATGCGTTCAAGTATGAAACGGCAA 693
QY 2055 GTACGACATTAAAGACGTGGCGTGGATACGCTGGCGGAAAGCGGGTCTGACCTTCCT 2114
DB 692 GTACGACATTAAAGACGTGGCGTGGATACGCTGGCGGAAAGCGGGTCTGACCTTCCT 633
QY 2115 GGTGACCTGATTAAACAAACACATGAATGACAGACCGGATTAATCTCCATCGAGAGC 2174
DB 632 GGTGACCTGATTAAACAAACACATGAATGACAGACCGGATTAATCTCCATCGAGAGC 573
QY 2175 TGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGTGGCGATGTTCCAAAT 2234
DB 572 TGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGTGGCGATGTTCCAAAT 513
QY 2235 CGACACAGCAAGTGAATTAATGTTGTTAAAGGTTACGCTGCGACCTTCAAGGTTCAACATC 2294
DB 512 CGACACAGCAAGTGAATTAATGTTGTTAAAGGTTACGCTGCGACCTTCAAGGTTCAACATC 453
QY 2295 CAACACCTGCTGTTGGCGTGTGAGCGCAGGTATTAACGGCGCAGTCCGAAACAGAGCT 2354
DB 452 CAACACCTGCTGTTGGCGTGTGAGCGCAGGTATTAACGGCGCAGTCCGAAACAGAGCT 393
QY 2355 GCGAAAGAGTTTCTCGAAACCTATCTGCTGATGATGAAGGTCTGGAAGCGGTTAA-TA 2413
DB 392 GCGAAAGAGTTTCTCGAAACCTATCTGCTGATGATGAAGGTCTGGAAGCGGTTAATTA 333

RESULT 14
AY080094/c
LOCUS
DEFINITION
  AY080094 Scripps Pier (La Jolla, CA) uncultured virus community
  uncultured marine virus genomic clone SIO51p3D4L, genomic survey
  sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  AY080094.1 GI:24745276
  GSS.
  uncultured marine virus
  uncultured marine virus
  Viruses; environmental samples.
REFERENCE
  1 (bases 1 to 608)
  Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M.,
  Mead, D., Azam, F. and Rohwer, F.
  Genomic analysis of uncultured marine viral communities
  Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
  22294988
  MEDLINE
  PUBMED
  12384570
  COMMENT
  Contact: Rohwer F
  Biology Dept.
  San Diego State University
  5500 Campanile Dr, San Diego, CA 92102, USA
  Tel: 6195941336
  Fax: 6195956676
  Email: forest@unstroke.sdsu.edu
  Class: shotgun.
  Location/Qualifiers
    1..608
    /organism="uncultured marine virus"
    /mol_type="genomic DNA"
    /db_xref="taxon:186617"
    /clone="SIO51p3D4L"
    /clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
    community"
    /note="Marine viruses were isolated from 200 liters of
    surface seawater using a combination of differential
    filtration and density-dependent gradient centrifugation.
    Linker-amplified shotgun libraries were constructed by
    randomly shearing the total marine viral community DNA,
    end-repairing, ligating dsDNA linkers to the ends, and
    amplifying the fragments using Vent DNA polymerase. The
    resulting fragments were ligated into the pSMART vector
    and electroporated into MC12 cells (Lucigen; Middleton,
    WI)."

ORIGIN
  source
  Query Match 17.9%; Score 589.8; DB 8; Length 608;
  Best Local Similarity 99.7%; Pred. No. 3.7e-163;
  Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 287 CAGTCTGGCCCTGCACGGCGCGTCCCAATTCGCGGCGATTAAATCTCGCGCGATCA 346
DB 608 CAGTCTGGCCCTGCACGGCGCGTCCCAATTCGCGGCGATTAAATCTCGCGCGATCA 549

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds
(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-15
Perfect score: 3300
Sequence: 1 ccgacacatcgaaatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues
Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	10	US-09-765-555-14
2	3300	100.0	3300	10	US-09-765-555-15
3	3226.4	97.8	3300	10	US-09-765-555-18
4	3215.2	97.4	3300	10	US-09-765-555-17
5	3207.2	97.2	3300	10	US-09-765-555-16
6	2705.8	82.0	7373	18	US-10-257-384A-1
7	2705.8	82.0	8101	17	US-10-149-472-5

8	2701	81.8	6648	19	US-10-343-859-8	Sequence 8, Appli
9	2698	81.8	9191	19	US-10-343-859-9	Sequence 9, Appli
10	2690.4	81.5	7112	18	US-10-263-153-40	Sequence 40, Appl
11	2690.4	81.5	7259	18	US-10-263-153-30	Sequence 30, Appl
12	2690.4	81.5	7322	18	US-10-263-153-35	Sequence 35, Appl
13	2690.4	81.5	7352	18	US-10-263-153-25	Sequence 25, Appl
14	2690.4	81.5	7370	18	US-10-263-153-20	Sequence 20, Appl
15	2690.4	81.5	7370	18	US-10-263-153-61	Sequence 61, Appl
16	2690.4	81.5	7370	18	US-10-263-153-66	Sequence 66, Appl
17	2690.4	81.5	7370	18	US-10-263-153-71	Sequence 71, Appl
18	2690.4	81.5	7403	18	US-10-263-153-15	Sequence 15, Appl
19	2690.4	81.5	7442	18	US-10-263-153-10	Sequence 10, Appl
20	2690.4	81.5	7478	18	US-10-263-153-3	Sequence 3, Appli
21	2636.8	79.9	6806	10	US-09-908-943A-194	Sequence 194, App
22	2636.8	79.9	6806	20	US-10-801-487-194	Sequence 194, App
23	2636.8	79.9	6806	20	US-10-801-938-194	Sequence 194, App
24	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
25	2636.8	79.9	6806	20	US-10-801-486-194	Sequence 194, App
26	2636.8	79.9	6806	21	US-10-801-493-194	Sequence 194, App
27	2605.4	48.8	7553	18	US-10-263-153-7	Sequence 7, Appli
28	1609.4	44.7	4700	17	US-10-241-596-137	Sequence 137, App
29	1474	44.7	4700	17	US-10-313-963A-54	Sequence 54, Appl
30	1462.4	44.3	4921	21	US-10-495-491-2	Sequence 2, Appli
31	1462.4	44.3	4935	21	US-10-495-491-1	Sequence 1, Appli
32	1462.4	44.3	4945	21	US-10-495-491-3	Sequence 3, Appli
33	1462.4	44.3	4951	21	US-10-495-491-4	Sequence 4, Appli
34	1417.6	43.0	1511	10	US-09-987-763-30	Sequence 30, Appl
C 35	1417.6	43.0	5926	13	US-10-024-809-3	Sequence 3, Appli
C 36	1238.4	37.5	5024	17	US-10-313-963A-52	Sequence 52, Appl
C 37	1238.4	37.5	5448	21	US-10-671-995A-2	Sequence 2, Appli
C 38	1238.4	37.5	5454	21	US-10-671-995A-1	Sequence 1, Appli
39	1237.2	37.5	1922	10	US-09-995-898A-32	Sequence 32, Appl
40	1237.2	37.5	1922	17	US-10-420-034A-32	Sequence 32, Appl
41	1237.2	37.5	6270	21	US-10-855-897-2	Sequence 2, Appli
42	1204	36.5	4557	17	US-10-457-372-5	GENERAL INFORMA
43	1204	36.5	6664	18	US-10-270-176-1	Sequence 1, Appli
44	1204	36.5	6668	18	US-10-270-176-2	Sequence 2, Appli
45	1204	36.5	6742	18	US-10-270-176-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-765-555-14
; Sequence 14, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765.555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m1 and zinc finger
; OTHER INFORMATION: protein ZfPm1
US-09-765-555-14

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTCAAAA	60
Db	1	CCGACACCATCGAATGGTCAAAA	60
QY	61	GTCAAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATTCGCAGAGTATGCGG	120
Db	61	GTCAAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATTCGCAGAGTATGCGG	120
QY	121	GTGTCTCTTATCAGACCCGTTTCCCGCTGGTGAACACAGCCAGCCACGTTTCTCGGAAAA	180
Db	121	GTGTCTCTTATCAGACCCGTTTCCCGCTGGTGAACACAGCCAGCCACGTTTCTCGGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTCAATTTACATTTCCCAACCGCGTGGCAC	240
Db	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTCAATTTACATTTCCCAACCGCGTGGCAC	240
QY	241	AACAACTGCGGGCAAAACAGTGTGCTGATGGGGTTTCCACCTCCAGTTCGGCCCTGC	300
Db	241	AACAACTGCGGGCAAAACAGTGTGCTGATGGGGTTTCCACCTCCAGTTCGGCCCTGC	300
QY	301	ACGCGCCGTGCGCAATTGTTCGGCGGATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCCGTGCGCAATTGTTCGGCGGATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG	360
QY	361	TGGTGGTTCGATGCGTAAGCGAGCGCGTGCAGAGCCTGTAAAGCGCGGTGCACAATC	420
Db	361	TGGTGGTTCGATGCGTAAGCGAGCGCGTGCAGAGCCTGTAAAGCGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG	480
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG	480
QY	481	TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA	540
Db	481	TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGCATCGGCGGTGAGCATC	600
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGCATCGGCGGTGAGCATC	600
QY	601	TGGTGCATTTGGGTCAACAGCAATTCGCGCTGTTAGCGGGCCCATTAAGTTCTCTCTCGG	660
Db	601	TGGTGCATTTGGGTCAACAGCAATTCGCGCTGTTAGCGGGCCCATTAAGTTCTCTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATCTCACTTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATCTCACTTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGGTTTCAACCAACCATGCAATGCTGA	780
Db	721	CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGGTTTCAACCAACCATGCAATGCTGA	780
QY	781	ATGAGGGATGTTTCCCACTGCGATGCTGGTGCACACGATCAGATGGCGCTGGGCGCAA	840
Db	781	ATGAGGGATGTTTCCCACTGCGATGCTGGTGCACACGATCAGATGGCGCTGGGCGCAA	840
QY	841	TGCGCGCATTAACGAGTCCGGCTGCGGCTGGTGGTGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTAACGAGTCCGGCTGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCCACCATCAAAACAGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCCACCATCAAAACAGATTTTC	960
QY	961	GCTGTGGGGCAAAACAGCGTGAACCGTTGTCGCAACTCTTCAGGGCCAGCGGTGA	1020
Db	961	GCTGTGGGGCAAAACAGCGTGAACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACACCTGGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACACCTGGCGCCCAATA	1080
QY	1081	CGCAAAACCGCCTCTCTCCCGCGCGTGGCCGATTCATTAATGCAGTGGCACGACAGGTTT	1140

Db	1081	CGCAAAACCGCCTCTCTCCCGCGCGTGGCCGATTCATTAATGCAGTGGCACGACAGGTTT	1140
QY	1141	CCCAGCTGGAAAGCGGCGAGTGAGCGCAACGCAATAATGTAGTTAGCTCAGTCATTAG	1200
Db	1141	CCCAGCTGGAAAGCGGCGAGTGAGCGCAACGCAATAATGTAGTTAGCTCAGTCATTAG	1200
QY	1201	GCACAAATTCATGCTTTTGACAGCTTATCATCGATGCACTGACCGGTGCACCAATGCTTCGGCG	1260
Db	1201	GCACAAATTCATGCTTTTGACAGCTTATCATCGATGCACTGACCGGTGCACCAATGCTTCGGCG	1260
QY	1261	TCAGGACAGCCATCGAAAGCTGTGGTATGGCTGTGCAGTCTGTAATCACTGCAATAATTCG	1320
Db	1261	TCAGGACAGCCATCGAAAGCTGTGGTATGGCTGTGCAGTCTGTAATCACTGCAATAATTCG	1320
QY	1321	TGTCGCTCAAGGGCGCACTCCCGTTCTGGATAATGTTTTTTTCGCGCGACATCAATAACGGTT	1380
Db	1321	TGTCGCTCAAGGGCGCACTCCCGTTCTGGATAATGTTTTTTTCGCGCGACATCAATAACGGTT	1380
QY	1381	CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATTAATGTGTGGA	1440
Db	1381	CTGGCAATATCTGAAATGAGCTGTGTGACAAATTAATCATCGGCTCGTATTAATGTGTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAAAGTGTTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAAAGTGTTTTTCACGA	1500
QY	1501	GCATTTCCACCAACAGGACCATAGATTATGAAAATCTGAAGAGTAAATCTGGTAATCTCG	1560
Db	1501	GCATTTCCACCAACAGGACCATAGATTATGAAAATCTGAAGAGTAAATCTGGTAATCTCG	1560
QY	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGGTCAACCGTTGAGCATTCGGATAAACTTGAAGAGAAATTCGCCACAGGTT	1680
Db	1621	ACCGGAATTAAGGTCAACCGTTGAGCATTCGGATAAACTTGAAGAGAAATTCGCCACAGGTT	1680
QY	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCCTGTGGCTGAAATCACCCCGCAAAAGCGTTCCAGGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCCTGTGGCTGAAATCACCCCGCAAAAGCGTTCCAGGACAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGCGGTAACGTTACACGGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
Db	1801	CGGTTTACCTGGGATGCGGTAACGTTACACGGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
QY	1861	GAAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAAACCCGCAAAACCTCGGAA	1920
Db	1861	GAAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAAACCCGCAAAACCTCGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACGTTGGGCGTGATTAACGCTGGCGGAAGCG	2100
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACGTTGGGCGTGATTAACGCTGGCGGAAGCG	2100
QY	2101	GGTCTGACCTTCTGCTGCTGATTAATAAACAACATCAATGACAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGCTGCTGATTAATAAACAACATCAATGACAGACACCGATTAC	2160
QY	2161	TCATCGCAGAGCTGCCCTTTTAAAGCGGAAACAGCGATGACCATCAACCGCCGCTGG	2220

Db 2161 TCCATCGCAGAGCTGCCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCAACATCGACACACGAAAGTGAATTATGGTGAACGGTACTGCGCACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGAAAGTGAATTATGGTGAACGGTACTGCGCACCTTC 2280
Qy 2281 AAGGTCACACATCAAAACCGTTCTGGCGTGTGAGCGGAGGTATTAAACGCCCGCAGT 2340
Db 2281 AAGGTCACACATCAAAACCGTTCTGGCGTGTGAGCGGAGGTATTAAACGCCCGCAGT 2340
Qy 2341 CGAAACAAAGAGCTGGCAAAAGAGTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CGAAACAAAGAGCTGGCAAAAGAGTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTAAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGAAAG 2460
Db 2401 GAAGCGGTAAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGAAAG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATTCGCCGCCACCATGGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTCGCCGCCACCATGGAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CGAAACATCCGACATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CGAAACATCCGACATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACACAAACATTAACAAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACAAACATTAACAAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCTTCTCTGTCGCGCAGCGGCCCTCGAGCCCGGGGAGAAGCCCTATGCTGT 2760
Db 2701 GGATCCTCTTCTCTGTCGCGCAGCGGCCCTCGAGCCCGGGGAGAAGCCCTATGCTGT 2760
Qy 2761 CGGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCACACAGCGTACCCAC 2820
Db 2761 CGGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCACACAGCGTACCCAC 2820
Qy 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTAGCGAGTCCAGCAAC 2880
Db 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTAGCGAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATCTGCGAGAGAGCCATACAAATGTCAGAAATGTGCG 2940
Db 2881 CTGGTGGCCATCAACGCACTCATCTGCGAGAGAGCCATACAAATGTCAGAAATGTGCG 2940
Qy 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCGGGCAACCAACGTAATCAACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCGGGCAACCAACGTAATCAACCGGGGAGAG 3000
Qy 3001 CCTATGCTGTGCGGAATGTGGTAAGTCTTTAGCTGAGCGGAGGATAAATCTGGTGGCCAC 3060
Db 3001 CCTATGCTGTGCGGAATGTGGTAAGTCTTTAGCTGAGCGGAGGATAAATCTGGTGGCCAC 3060
Qy 3061 GAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTAGC 3120
Db 3061 GAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTAGC 3120
Qy 3121 CAGGCGGGCACTGGCCAGCCATCAACGCACTCATCTGCGGAGAGGAGCATACAAATGT 3180
Db 3121 CAGGCGGGCACTGGCCAGCCATCAACGCACTCATCTGCGGAGAGGAGCATACAAATGT 3180
Qy 3181 CCAGAAATGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGCAACCAACGTAATCAAC 3240
Db 3181 CCAGAAATGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGCAACCAACGTAATCAAC 3240
Qy 3241 ACCGGTAAAAAAACTAGTGGCCAGGCGCGCGAGTACCGGTACGAGTTCGGAGCTACGCT 3300
Db 3241 ACCGGTAAAAAAACTAGTGGCCAGGCGCGCGAGTACCGGTACGAGTTCGGAGCTACGCT 3300

RESULT 2

US-09-765-555-15
; Sequence 15, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger
; OTHER INFORMATION: protein ZfPm2
US-09-765-555-15

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGACACCATCAATGGTGC AAAACCTTTTCGCGGTATGGCATGATAGCCCGGGAAGAGA 60
Db 1 CCGACACCATCAATGGTGC AAAACCTTTTCGCGGTATGGCATGATAGCCCGGGAAGAGA 60
Qy 61 GTCAAATTCAGGTCGTGAATGTGAAAACCAAGTAAAGTTATACAGATGTCGAGAGTATGCG 120
Db 61 GTCAAATTCAGGTCGTGAATGTGAAAACCAAGTAAAGTTATACAGATGTCGAGAGTATGCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAAACCAAGCCAGCCAGCCAGTCTTCGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAAACCAAGCCAGCCAGCCAGTCTTCGCGAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGAGTGAATTAATCCCAACCCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGAGTGAATTAATCCCAACCCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAACAGTCGTTGATGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGGCAACAGTCGTTGATGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Qy 301 ACGGCGCTGCGCAAAATTCGCGCGGATTAATCTCGCGCGATCAACTCGGCGGATCACTGGGTGCGAGG 360
Db 301 ACGGCGCTGCGCAAAATTCGCGCGGATTAATCTCGCGCGGATCAACTCGGCGGATCACTGGGTGCGAGG 360
Qy 361 TGGTGGTTCGATGGTAGAAGCAAGCGCGCTCGAAGCCCTGTAAGCGGGGTGCACAATC 420
Db 361 TGGTGGTTCGATGGTAGAAGCAAGCGCGCTCGAAGCCCTGTAAGCGGGGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCGGTGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCGGTGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTCCGCAATGGGTCCACAGCAAAATTCGCGCTGTTAGCGGGGCCCAATTAAGTTCTGTCGG 660

Db 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATAAGATTCTGTCTCGG 660
Qy 661 CGCGTCTGCGTCTGCTGCGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCTGCGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAAACGGAAAGCGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAAACGGAAAGCGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTGGATGCTGGTTGCCAAGATCGATGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCGTTCCCACTGGATGCTGGTTGCCAAGATCGATGCGCTGGCGCGAA 840
Qy 841 TGCGGCCCATTAACCAAGTCCGGCTGCGGTGGTGGCGGATCTCTCGGTAGTGGGATACG 900
Db 841 TGCGGCCCATTAACCAAGTCCGGCTGCGGTGGTGGCGGATCTCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAACATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAACATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTCTGTCGAATCTCTCAGGGCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTCTGTCGAATCTCTCAGGGCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGTGGAAGAAACCAACCTGCGGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGTGGAAGAAACCAACCTGCGGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCAATTAATGCAGTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCAATTAATGCAGTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGGCAACCGCAATTAATGTAGTTAGTCACTCATATAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGGCAACCGCAATTAATGTAGTTAGTCACTCATATAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGACCAATGCTTCTCGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGACCAATGCTTCTCGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGTATGGTGTGCGAGTGTAAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGTATGGTGTGCGAGTGTAAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACCGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACCGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Qy 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAAACAGCCAGTCCGTTAGGTGTTTTACGA 1500
Db 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAAACAGCCAGTCCGTTAGGTGTTTTACGA 1500
Qy 1501 GCATCTCACCAACAGGACCATAGATTATGAAATCTGAAGAGTAACTGGTAAATCTGG 1560
Db 1501 GCATCTCACCAACAGGACCATAGATTATGAAATCTGAAGAGTAACTGGTAAATCTGG 1560
Qy 1561 ATTAACCGCGATTAAGGCTATAACCGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATTAAGGCTATAACCGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATAAATCGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATAAATCGGAAGAGAAATTTCCACAGGTT 1680
Qy 1681 CGCGCAATCGCGATGCGCCTGCATTTATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740

Db 1681 CGCGCAATCGCGATGCGCCTGCATTTATCTCTGGGCACACGACCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGTGTGAATCAACCCCGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGTGTGAATCAACCCCGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTCAACGGCAAGCTGATTTGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTCAACGGCAAGCTGATTTGCTTACCGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGGCTGGATTAAGAACTGAAGAGGAAAGGTAAAGAGCGCGCTGTATTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATTAAGAACTGAAGAGGAAAGGTAAAGAGCGCGCTGTATTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTTGCTGACCGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTTGCTGACCGGGGTTATGCTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTTAAAGAGCTGGCGTGGATAAACGCTGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGAGCTGGCGTGGATAAACGCTGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAAACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAAGCTGCTTTTAATAAGGGAACAGCATGACCATCAACGCCCGGTGG 2220
Db 2161 TCCATCGCAAGCTGCTTTTAATAAGGGAACAGCATGACCATCAACGCCCGGTGG 2220
Qy 2221 GCATGTCCAACATCAACACCGTAAAGTGAATTTATGTTGTAAACGCTACTGCCGACCTTC 2280
Db 2221 GCATGTCCAACATCAACACCGTAAAGTGAATTTATGTTGTAAACGCTACTGCCGACCTTC 2280
Qy 2281 AAGGTCACACCATCCAAACCGTTCTGTTGGGTGCTGAGCGCAGGTATTAAACGCCGCGAGT 2340
Db 2281 AAGGTCACACCATCCAAACCGTTCTGTTGGGTGCTGAGCGCAGGTATTAAACGCCGCGAGT 2340
Qy 2341 CCGAAACAAAGAGCTGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTCGCGCCACCATGGAAGAGCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTCGCGCCACCATGGAAGAGCGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGGTATGCGCTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGGTATGCGCTGCTACTGCGGTGATCAACGCC 2580
Qy 2581 GCGAGCGGTGTCAGACTGTGATGAAGCCCTTGAAGAGCGCGACGACTAATTCGAGTCTG 2640
Db 2581 GCGAGCGGTGTCAGACTGTGATGAAGCCCTTGAAGAGCGCGACGACTAATTCGAGTCTG 2640
Qy 2641 AACACAAACAAATAACAAATTAACACACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATAACAAATTAACACACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCTTCTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAAAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAAAGCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCAGCTGGTGGCCACAGGATCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCAGCTGGTGGCCACAGGATCCAC 2820

```
QY 2821 ACCGCTGAACACCGTATATAATGCCAGAGTGCAGCAAACTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACCGCTGAACACCGTATATAATGCCAGAGTGCAGCAAACTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCACTCAACGCACTCATACTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCCACTCAACGCACTCATACTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
QY 2941 AAGTCTTTTCTCTCGGTCTGACAACTCTCGTCCGGCACCAACGTACTACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTTCTCTCGGTCTGACAACTCTCGTCCGGCACCAACGTACTACACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTCTCAGCCGACGGGATAAAGCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTCTCAGCCGACGGGATAAAGCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATATAATGCCAGAGTGGCGGCAAACTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATATAATGCCAGAGTGGCGGCAAACTTTTAGC 3120
QY 3121 CAGCGCGGCCACTTGGCCAGCCATCAACGCACTCATACTGGCGAGAAGCCATACAAATGT 3180
Db 3121 CAGCGCGGCCACTTGGCCAGCCATCAACGCACTCATACTGGCGAGAAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTTCTCGGTCTGACAACTCTGCTCCGGCACCAAGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTTCTCGGTCTGACAACTCTGCTCCGGCACCAAGTACTCAC 3240
QY 3241 ACCGCTAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTAAGCTTCCGAGCTACGCT 3300
Db 3241 ACCGCTAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTAAGCTTCCGAGCTACGCT 3300
```

RESULT 3

```
US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-Ap3 and zinc finger
; OTHER INFORMATION: protein ZFPap3
US-09-765-555-18
```

```
Query Match 97.8%; Score 3226.4; DB 10; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
```

```
QY 1 CCGACACCATCAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGA 60
Db 1 CCGACACCATCAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGA 60
QY 61 GTCAATTTCAGGGTGGTGAATGTCAAAACCAAGTAACTGTAACGATGTCGAGAGTATCGCG 120
Db 61 GTCAATTTCAGGGTGGTGAATGTCAAAACCAAGTAACTGTAACGATGTCGAGAGTATCGCG 120
QY 121 GTGTCCTTTATCAGACCGGTTTCCCGGTGGTGAACCAAGCCAGCCACGCTTCTCGCAAAA 180
```

```
Db 121 GTGTCCTTTATCAGACCGGTTTCCCGGTGGTGAACCAAGCCAGCCACGCTTCTCGCAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGTGATTTGGGCTTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTGATTTGGGCTTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCGCTCGCAAAATGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCGCTCGCAAAATGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 361 TGGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATCTATCCGCTGGATGACCAAGTGGCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATCTATCCGCTGGATGACCAAGTGGCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGGCTTATTTCTTGATGTCCTGACCA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGGCTTATTTCTTGATGTCCTGACCA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGTACGCGACTGGCGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGTACGCGACTGGCGCTGGAGCATC 600
QY 601 TGGTGCATTTGGGTACACAGCAAAATCGGCTGTTAGCGGGCCCATTAAGTTCGTCTCG 660
Db 601 TGGTGCATTTGGGTACACAGCAAAATCGGCTGTTAGCGGGCCCATTAAGTTCGTCTCG 660
QY 661 CGGCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAAGGCGACTGGAGTGCCATGTCGGTTCCTCAACAAACCAATGCAATGCTGA 780
Db 721 CGGAAACGGGAAAGGCGACTGGAGTGCCATGTCGGTTCCTCAACAAACCAATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCATCGCATGCTGGTGTGCAACAGATCAGATGCGCTGGCGCA 840
Db 781 ATGAGGGCATCGTTCCCATCGCATGCTGGTGTGCAACAGATCAGATGCGCTGGCGCA 840
QY 841 TGGCGCCATTACCGAGTCCCGGCTGCGGCTGCGGATATCTCGTGTGAGTGGGATAG 900
Db 841 TGGCGCCATTACCGAGTCCCGGCTGCGGCTGCGGATATCTCGTGTGAGTGGGATAG 900
QY 901 ACGATACCGAAAGACAGCTCATGTTTATATCCCGCGCTTAAACCAACCAATCAACAGGATTTTC 960
Db 901 ACGATACCGAAAGACAGCTCATGTTTATATCCCGCGCTTAAACCAACCAATCAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAAAACCAAGCGTGGACCGCTTGTGTCGAATCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACCAAGCGTGGACCGCTTGTGTCGAATCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGTGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGTGAAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGCAGTGGCACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGCAGTGGCACGACAGGTTT 1140
QY 1141 CCGCATGGAAGCGGGCAGTGGAGCGAAACGCAATTAATGTAGTGTAGTCACTCATTTAG 1200
Db 1141 CCGCATGGAAGCGGGCAGTGGAGCGAAACGCAATTAATGTAGTGTAGTCACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTTCAGAGCTTATCATCGACTGCGAGGTGCAACCAATGCTTCGCGG 1260
```

Db 1201 GCACAAATCTCTATGTTTGACAGCTTATCATCGACTGACGCGTGCCACCAATGTTCTTGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGGTATGCTGTGAGGTGCTGTAATCACTGCATAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGCTGTGAGGTGCTGTAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACCTCCGTTCTGGATTAATGTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCGTTCTGGATTAATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTGTAATGTTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTGTAATGTTGGA 1440
QY 1441 ATTGTAGCGGATACAAATTTTACACAGAGAAACAGCCAGTCCGTTTAGGTGTTTTTACGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTACACAGAGAAACAGCCAGTCCGTTTAGGTGTTTTTACGA 1500
QY 1501 GCATTTCACCAACAGGACCATAGATTATGAAACTGGAAGGTAAACTGGTAATCTGG 1560
Db 1501 GCATTTCACCAACAGGACCATAGATTATGAAACTGGAAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAGTAAAGTAAAGTAAACTGGTAATCTGG 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAGTAAAGTAAAGTAAACTGGTAATCTGG 1620
QY 1621 ACCGGAATTAAGTACACGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGTT 1680
Db 1621 ACCGGAATTAAGTACACGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGTT 1680
QY 1681 GCGCACTGCGGATGGCCCTGACATTAATCTTTGCGGACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCACTGCGGATGGCCCTGACATTAATCTTTGCGGACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAATACACCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAATACACCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGGAACCCGCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGGAACCCGCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGCGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGCGAGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGATTAAGACGTTGGCGTGGATACGCTGCGGGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGATTAAGACGTTGGCGTGGATACGCTGCGGGAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAACAAACACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAACAAACACATGAATGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTATTAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTATTAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
QY 2221 GCATGTTCCAACTACGACACGACAAAGTGAATGATGCTGTAACGCTGCTGCGACCTTC 2280
Db 2221 GCATGTTCCAACTACGACACGACAAAGTGAATGATGCTGTAACGCTGCTGCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCGAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCGAGT 2340

RESULT 4

US-09-765-555-17
; Sequence 17, Application US/09765555
; Publication No. US2003003735A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate

QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTAATTGCGCGCACCATGGAACAAACCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTAATTGCGCGCACCATGGAACAAACCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGATGATCGGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGATGATCGGTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGTCAGACTGTCGATGAGCCCTCAAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGTCAGACTGTCGATGAGCCCTCAAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAAATAACAACAACTCGGGATCGAGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAAATAACAACAACTCGGGATCGAGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTTGTTAAGTCTTCTCTCAGAGCTCTCACTGTTGCGGCCACACAGCTGACCCAC 2820
Db 2761 CCGGAATGTTGTTAAGTCTTCTCAGCAGAGCAGCTCCTCTGTTGCGGCCACACAGCTGACCCAC 2820
QY 2821 ACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGGCAAAATCTTTTACGCCAGTCCAGCAAC 2880
Db 2821 ACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGGCAAAATCTTTTACGCCAGTCCAGCAAC 2880
QY 2881 CTGTTGCGGCATCAACGCACTCATCTGCGAGAGCCATACAAATGTCCAGAAATCTGGC 2940
Db 2881 CTGTTGCGGCATCAACGCACTCATCTGCGAGAGCCATACAAATGTCCAGAAATCTGGC 2940
QY 2941 AAGTCTTCTCTCGGTCTGACAAATCTCTCGCGCACCAACAGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTCTCAGCCAGTCCAGCAACCTGTTGGTGGCCACCAACAGTACTCACACCGGGGAGAAG 3000
QY 3001 CCTATGCTGTGTCGGAATGTGTGTAAGTCTTTCAGCGCGCAGCGATAACTGTTGCGCCAC 3060
Db 3001 CCTATGCTGTGTCGGAATGTGTGTAAGTCTTTCAGCACCAAGTGGCTCTTGGTTAGACAC 3060
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGCGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
QY 3121 CAGCGCGGCACTCTGCGCAGCCATCAACGCACTCATACTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGGCACTCTGCGCAGCCATCAACGCACTCATACTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTCTCTCGTCTGCAATCTCTGCGCGCACCAACGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTCTCGTCTGCAATCTCTGCGCGCACCAACGTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTAGCAGCTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTAGCAGCTTCCGGACTACGCT 3300

;
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein ZFPm4
US-09-765-555-17

Query Match 97.4%; Score 3215.2; DB 10; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 CGGACACCATCGAATGGTGCAGAACTTTTCGGGGTATGCGATGAGCGCCGGGAAGA 60
DB |||||
QY 1 CGGACACCATCGAATGGTGCAGAACTTTTCGGGGTATGCGATGAGCGCCGGGAAGA 60
DB |||||
QY 61 GTCAATTACAGGGTGTGAATGTGAACACCAAGTAACGTTATACGATGTCGAGAGTATGCGG 120
DB |||||
QY 61 GTCAATTACAGGGTGTGAATGTGAACACCAAGTAACGTTATACGATGTCGAGAGTATGCGG 120
DB |||||
QY 121 GTGTCTTATCAGACCGTTTCCCGGTGTGAACACGAGCCAGCCAGCGTTTCTGCGAAAA 180
DB |||||
QY 121 GTGTCTTATCAGACCGTTTCCCGGTGTGAACACGAGCCAGCCAGCGTTTCTGCGAAAA 180
DB |||||
QY 181 CGCGGAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCGCCAACCGCGTGGCAC 240
DB |||||
QY 181 CGCGGAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCGCCAACCGCGTGGCAC 240
DB |||||
QY 241 AACAACTGGCGGCAACAGTGTGCTGATTTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
DB |||||
QY 241 AACAACTGGCGGCAACAGTGTGCTGATTTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
DB |||||
QY 301 AGCGCGGTGCGCAAAATTTGCGGCGAATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
DB |||||
QY 301 AGCGCGGTGCGCAAAATTTGCGGCGAATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
DB |||||
QY 361 TGGTGTGTCGATGTAGAACGAGCGGCGTGAAGCGGCTGTAAAGCGGCGGTGCACAATC 420
DB |||||
QY 361 TGGTGTGTCGATGTAGAACGAGCGGCGTGAAGCGGCTGTAAAGCGGCGGTGCACAATC 420
DB |||||
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
DB |||||
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
DB |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTTCGGCGGTATTTCTTGATGTCCTGACACAGA 540
DB |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTTCGGCGGTATTTCTTGATGTCCTGACACAGA 540
DB |||||
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
DB |||||
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
DB |||||
QY 601 TGGTGTGATTTGGTTCACAGCAAACTCGGCTGTAGCGGGCCCATTAAGTTCTGTCCTGG 660
DB |||||
QY 601 TGGTGTGATTTGGTTCACAGCAAACTCGGCTGTAGCGGGCCCATTAAGTTCTGTCCTGG 660
DB |||||
QY 661 CGCGTCTGCGTCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB |||||
QY 661 CGCGTCTGCGTCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB |||||
QY 721 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGTTTCAACAAACCATGCAAAATGCTGA 780
DB |||||

DB 721 CGGAACGGGAAGCGGACCTGGAGTGCCATGTCGGTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCAACGATCAGATGCGTGGCGCGAA 840
DB |||||
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCAACGATCAGATGCGTGGCGCGAA 840
DB |||||
QY 841 TGGCGGCATTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGTAGTGGGATAG 900
DB |||||
QY 841 TGGCGGCATTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGTAGTGGGATAG 900
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACGAGTTTC 960
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACGAGTTTC 960
DB |||||
QY 961 GCCTGCTGGGCAAAACCAAGCGTGCACCGTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
DB |||||
QY 961 GCCTGCTGGGCAAAACCAAGCGTGCACCGTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
DB |||||
QY 1021 AGGCAATCAGCTGTGCCCCGTCTCATCTGGTGAAGAAAGAAACCAACCCCTGGCGCCCAATA 1080
DB |||||
QY 1021 AGGCAATCAGCTGTGCCCCGTCTCATCTGGTGAAGAAAGAAACCAACCCCTGGCGCCCAATA 1080
DB |||||
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGATTCAATTAATCAGCTGCGACGACGAGTTT 1140
DB |||||
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGATTCAATTAATCAGCTGCGACGACGAGTTT 1140
DB |||||
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
DB |||||
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCGTGCACCAATGCTTCTGGCG 1260
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCGTGCACCAATGCTTCTGGCG 1260
DB |||||
QY 1261 TCAGCAGCCATCGGAAGCTGTGGTGTGAGCTGTGAGCTGCTGTAATCACTGCATAATTCG 1320
DB |||||
QY 1261 TCAGCAGCCATCGGAAGCTGTGGTGTGAGCTGTGAGCTGCTGTAATCACTGCATAATTCG 1320
DB |||||
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTGCGCGGACATCAATACGGTT 1380
DB |||||
QY 1381 CTGGCAATTTCTGAAATGAGCTGTGCAATTAATCATCGGCTCGTATATGTTGTGA 1440
DB |||||
QY 1381 CTGGCAATTTCTGAAATGAGCTGTGCAATTAATCATCGGCTCGTATATGTTGTGA 1440
DB |||||
QY 1441 ATTTGAGCGGATAACAAATTTCAACAGGAAACAGCAGTCCGTTTAGTGTTCACGA 1500
DB |||||
QY 1441 ATTTGAGCGGATAACAAATTTCAACAGGAAACAGCAGTCCGTTTAGTGTTCACGA 1500
DB |||||
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAAGAGGTAAACTGGTAACTCTGG 1560
DB |||||
QY 1561 ATTTAAGCGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1561 ATTTAAGCGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1621 ACCGGAATTAAGTACCGCTTGGATCCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
DB |||||
QY 1621 ACCGGAATTAAGTACCGCTTGGATCCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
DB |||||
QY 1681 CGCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGACACGACCGCTTTGGTGGCTAC 1740
DB |||||
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGACAAGCTGTAT 1800
DB |||||
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGACAAGCTGTAT 1800
DB |||||
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCGAGCTGATTTGCTTACCCGATCGCTGT 1860
DB |||||
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCGAGCTGATTTGCTTACCCGATCGCTGT 1860
DB |||||


```
QY 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGCTGACGGGGGTATTGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGCTGACGGGGGTATTGCGTTCAAG 2040
QY 2041 TATGAAAACCGGCAAGTACGACATTAAGACGTTGGCGCTGGATAACCGTGGCGGAAACGC 2100
Db 2041 TATGAAAACCGGCAAGTACGACATTAAGACGTTGGCGCTGGATAACCGTGGCGGAAACGC 2100
QY 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAATAAGCGGAACAGCGATGACCATCAACGGCCGCTGG 2160
Db 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAATAAGCGGAACAGCGATGACCATCAACGGCCGCTGG 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGGAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGGAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACACGAGCAAGTGAATTATGGTGTAAACGGTACTCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGAGCAAGTGAATTATGGTGTAAACGGTACTCCGACCTTC 2280
QY 2281 AAGGTCACCAATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACCTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACCTATCTGCTGACTGATGAAGGCTG 2400
QY 2401 GAAGCGTTTAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTTAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTGCGCCACCATGGAACCGCCAGAACCGCCGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTGCGCCACCATGGAACCGCCGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAATAACAATAACAACCTCTCGGATCGAGGGAAGATTTTCAGATTTC 2700
Db 2641 AACCAACAAACAATAACAATAACAACCTCTCGGATCGAGGGAAGATTTTCAGATTTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGCGCCCTTCGAGCGCCGGGAGAAAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGCCCTTCGAGCGCCGGGAGAAAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAGTCTTCTCTCAGAGCTCTCACTGTGTGCGCCACCGAGCGTACCAC 2820
Db 2761 CCGGAATGTGTAGTCTTCTCTCAGAGCTCTCACTGTGTGCGCCACCGAGCGTACCAC 2820
QY 2821 ACCGGTGAATAACCGTATAAATGCCAGAGTGGCGGCAAACTCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACCGGTGAATAACCGTATAAATGCCAGAGTGGCGGCAAACTCTTTAGCCAGAGCAGCAGC 2880
QY 2881 CTGTGTGCGCCATCAACGCATCATATCTGGCGAGAGGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGTGTGCGCCATCAACGCATCATATCTGGCGAGAGGCCATACAAATGTCCAGATGTGGC 2940
```

```
QY 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTACACCGGGGAGAAG 3000
QY 3001 CCTATGCTTGTTCGGAAATGTGGTAAAGTCTTTCAGCGCAGCGAATAACCTGGTGGCCAC 3060
Db 3001 CCTATGCTTGTTCGGAAATGTGGTAAAGTCTTTCAGCGCAGCGAATAACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCACACGCGGTGAAACACCGTATAAATGCCCAGAGTGGCGCAATCTTTTACG 3120
Db 3061 CAGCGTACCACACGCGGTGAAACACCGTATAAATGCCCAGAGTGGCGCAATCTTTTACG 3120
QY 3121 CAGCGCCGACCTGCGCAGCATCAACGCACTCATACTGGCGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCCGACCTGCGCAGCATCAACGCACTCATACTGGCGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTCAGGCGCATTTGGTCCGTCAACAAACGTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCGCATACCCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCGCATACCCGTACGACGTTCCGGACTACGCT 3300
```

RESULT 5

```
US-09-765-555-16
; Sequence 16, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m3 and zinc finger
; OTHER INFORMATION: protein ZFPm3
US-09-765-555-16
```

```
Query Match          97.2%; Score 3207.2; DB 10; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAAAACCCAGTAACTGATATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAAAACCCAGTAACTGATATACGATGTCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGAACCCAGCCAGCCAGCTTCTTCGCCGAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGAACCCAGCCAGCTTCTTCGCCGAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGGCGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGGCGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 ACAACATGGCGGGCAACACAGTGGTCTGATGGCGGTGCCACCTCAGTCTGGCCCTGCG 300
```


Db 241 AACAACTGGCGGGCAAA CAGTGGTCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTCGCAAAATFTGCGCGGATTAATCTCGGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTCGCAAAATFTGCGCGGATTAATCTCGGCGCGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGGTGTGATGTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGTGCACAATC 420
Db 361 TGGTGGTGTGATGTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGTGGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGAAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTCCTGACCAAGA 540
Db 481 TTGCTGTGAAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTCCTGACCAAGA 540
Qy 541 CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGGTCAACAGCAAACTCGGCTGTAGCGGGCCCATTTAAGTTCGTCTGG 660
Db 601 TGGTGGCATTTGGGTCAACAGCAAACTCGGCTGTAGCGGGCCCATTTAAGTTCGTCTGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCACGCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCACGCGATAG 720
Qy 721 CGGAACGGGAACGCACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAACGCACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATCTGCCACTGCGATGCTGTTGCCAAACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATCTGCCACTGCGATGCTGTTGCCAAACGATCAGATGGCGTGGCGCAA 840
Qy 841 TCGCGGCCATACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCCATACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
Qy 961 GCCTGTGGGGCAACACGCGTGGACCGCTTGTGCAACTCTCAGGSCCAGGCGGTGA 1020
Db 961 GCCTGTGGGGCAACACGCGTGGACCGCTTGTGCAACTCTCAGGSCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGGTTT 1140
Qy 1141 CCGGACTGGAAGCGGAGTGGCGCAACGCAATTAATGTCAGTTAGCTCACTCATTAG 1200
Db 1141 CCGGACTGGAAGCGGAGTGGCGCAACGCAATTAATGTCAGTTAGCTCACTCATTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATGCACTGCAGCGTGCAACCAATGCTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATGCACTGCAGCGTGCAACCAATGCTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAAGTCTGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAAGTCTGTAATCACTGCATTAATTCG 1320
Qy 1321 TGTGCTCAAGGGCACTCCGTTCTGGATATGTTTTTGGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGGGCACTCCGTTCTGGATATGTTTTTGGCGGACATCAATACGGTT 1380

Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCTGGCTCGTATTAATGTTGCGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCTGGCTCGTATTAATGTTGCGA 1440
Qy 1441 ATTGTAGCGGATACAAATTTTACACAGGAAACAGCAGTCCGTTTACGTTTACGTTTACGGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTACACAGGAAACAGCAGTCCGTTTACGTTTACGGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCTGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCTGTAATCTGG 1560
Qy 1561 ATTAAACGGCGATTAAGGCTATAAACGGTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGCGATTAAGGCTATAAACGGTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGGCTAGTTACAAACGGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCTAGTTACAAACGGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
Qy 1861 GAAAGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAAGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGCTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGCTGGATAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAAGACACCGATTAC 2160
Qy 2161 TCCATCGAAGAGCTGCTCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGAAGAGCTGCTCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCATCAAAACCGTTTCTGGCGTGTGAGCGCAAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTCACCATCAAAACCGTTTCTGGCGTGTGAGCGCAAGGTATTAAACGCCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGCTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGCTCTG 2400
Qy 2401 GAAAGCGTTTAAATAAGACAAACCGCTGGGTGCCGTAGCGTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAAGCGTTTAAATAAGACAAACCGCTGGGTGCCGTAGCGTGAAGTCTTTACGAGGAAG 2460

```
QY 2461 TTGGCGAAGATCCACGATTTCCGCCACCATCGGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGATTTCCGCCACCATCGGAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAATATCCCGAGATGTCGCGCTTTCTGGTATGCGGTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAATATCCCGAGATGTCGCGCTTTCTGGTATGCGGTGCTGCGGTGATCAACGCC 2580
QY 2581 GCACGGGTGTCAGATGTCGATGAAGCCCTGGAAGAGCGGCAGACTAATTCGAGCTCG 2640
Db 2581 GCACGGGTGTCAGATGTCGATGAAGCCCTGGAAGAGCGGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTCAGAAATC 2700
Db 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTCAGAAATC 2700
QY 2701 GGATCCTCTTCTCTGTCGCGCCAGCGGCCCTCGAGCCCGGGGAGAACCTATGCTTGT 2760
Db 2701 GGATCCTCTTCTCTGTCGCGCCAGCGGCCCTCGAGCCCGGGGAGAACCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCTGCTGCGCCACCGAGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTCTCAGAGCTCTCGCCACCTGCTGCGCCACCGAGTACCCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCGCAGAGTGCAGCAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCGCAGAGTGCAGCAATCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGCCCATCAACGCACTCATACTGCGGAGAGCCATACAATGTCAGATGTGGC 2940
Db 2881 CTGGTGCCCATCAACGCACTCATACTGCGGAGAGCCATACAATGTCAGATGTGGC 2940
QY 2941 AAGTCTTCTCTCGGTCTGCAATCTCTGTCGCGCAACCACTACTCACACCGGGAGAG 3000
Db 2941 AAGTCTTCTCAGCAGAGCTCAGCTGTCGCGCAACCACTACTCACACCGGGAGAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCAGCCAGAGCAGTCTCCCTGTCGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCAGCCAGAGCAGTCTCCCTGTCGCCAC 3060
QY 3061 CAGGTATCCACACGGGTGAAAAACCGTATAAATGCGCAGAGTGCAGCAATCTTTAGC 3120
Db 3061 CAGGTATCCACACGGGTGAAAAACCGTATAAATGCGCAGAGTGCAGCAATCTTTAGC 3120
QY 3121 CAGGC CGGCCACCTGCGCAGCCATCAACGCACTCATACTGCGGAGAGCCATACAATGT 3180
Db 3121 GACTGCGCGACCTTGCTCGCCATCAAGCACTCATACTGCGGAGAGCCATACAATGT 3180
QY 3181 CCAGATGTGCAAGTCTTTCTCTGCGTCTGACAACTCTGTCGCGCACCACAGTACTCAC 3240
Db 3181 CCAGATGTGCAAGTCTTTCTCTGCGTCTGACAACTCTGTCGCGCACCACAGTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCGACGCGCGGCGAGTACCGGTACGAGTTCGGGACTACGT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCGACGCGCGGCGAGTACCGGTACGAGTTCGGGACTACGT 3300

RESULT 6
US-10-257-384A-1
; Sequence 1, Application US/10257384A
; Publication No. US20040087524A1
; GENERAL INFORMATION:
; APPLICANT: Wiederanders, Bernd
; APPLICANT: Maubach, Gunter
; TITLE OF INVENTION: Agent for postoperative use after removal of bone tumors
; FILE REFERENCE: 2945-101
; CURRENT APPLICATION NUMBER: US/10/257,384A
; PRIOR FILING DATE: 2003-08-21
; PRIOR FILING DATE: 2001-04-18
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 19
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7373
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Expression vector pMalc2cchbmp2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)..(3453)
; OTHER INFORMATION: /product = fusion protein MBP Cystatin C BMP-2
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2689)
; OTHER INFORMATION: Fakter Xa Protease /position=388
; OTHER INFORMATION: (Amino acid sequence)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3067)..(3108)
; OTHER INFORMATION: /spacer between Cystatin C andBMP-2
; PUBLICATION INFORMATION:
; AUTHORS: Zwick, M B.
; TITLE: Expression vector pMal-X, complete sequence
; JOURNAL: Anal. Biochem.
; VOLUME: 264
; ISSUE: 1
; PAGES: 87-97
; DATE: 1998-11-01
; DATABASE accession NUMBER: Entrez Nucleotide database/AF031813
; DATABASE ENTRY DATE: 2001-05-07
; RELEVANT RESIDUES: 1528 - 3453
US-10-257-384A-1
```

```
Query Match 82.0%; Score 2705.8; DB 18; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCCGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCCGAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAAACCACTGTAACGTTATACGATGTCCAGAGATGTCGCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAAACCACTGTAACGTTATACGATGTCCAGAGATGTCGCG 120
QY 121 GTGCTCTTATCAGACGCTTTCCCGCGTGTGTAAACAGGCCAGCCAGCTTTCTTGCAGAAA 180
Db 121 GTGCTCTTATCAGACGCTTTCCCGCGTGTGTAAACAGGCCAGCCAGCTTTCTTGCAGAAA 180
QY 181 CCGCGGAAAAAAGTGGAAAGCGCGATGCGGAGCTGAAATTACATTCCCAACCGCGTGCAC 240
Db 181 CCGCGGAAAAAAGTGGAAAGCGCGATGCGGAGCTGAAATTACATTCCCAACCGCGTGCAC 240
QY 241 AACAACTGGCGGCAACAGTCTGCTGATTGGCGTTGCGCACCTCCAGTCTGCGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCTGCTGATTGGCGTTGCGCACCTCCAGTCTGCGCCCTGC 300
QY 301 ACGGCGCGTCCGAAATTTGTCGCGCGCATTAATCTCGCGCGCATCACTGGGTGCGCAGCG 360
Db 301 ACGGCGCGTCCGAAATTTGTCGCGCGCATTAATCTCGCGCGCATCACTGGGTGCGCAGCG 360
QY 361 TGGTGTGTGTCGATGGTAGAAACGAAAGCGCGTGAAGCCCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGGTGTGTGTCGATGGTAGAAACGAAAGCGCGTGAAGCCCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTCTGGAAGCTGCTGCACCTAATGTTCCGCGGTATTTCTTGATGCTCTTGACAGA 540
Db 481 TTGCTCTGGAAGCTGCTGCACCTAATGTTCCGCGGTATTTCTTGATGCTCTTGACAGA 540
```

QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGGACTGGCGTGGAGCATC 600
QY 601 TGGTCGATTTGGTTCACAGCAAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGATTTGGTTCACAGCAAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACACTGGAAGTGCATGTCGGTTTTTCAACAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACACTGGAAGTGCATGTCGGTTTTTCAACAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCACTGGATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCGTTCCTCACTGGATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGAA 840
QY 841 TGCGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGCGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGAGTGGAGCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGTGGAGCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCCCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCCCTGGCGGCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGTGGCAGCAGAGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAGCGTGCACCAATGCTTCTGGG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAGCGTGCACCAATGCTTCTGGG 1260
QY 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTTCGTAATCATCTGCATTAATTCG 1320
Db 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTTCGTAATCATCTGCATTAATTCG 1320
QY 1321 TGTCTCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACCGTT 1380
Db 1321 TGTCTCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACCGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA 1440
QY 1441 ATTGTAGGGGATAACAATTTTACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGGGGATAACAATTTTACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
QY 1501 GCACCTTCCACACAGGACCATAGATTTATGAAAAATCGAAGAGGTAAATCGTATCTGG 1560
Db 1501 GCACCTTCCACACAGGACCATAGATTTATGAAAAATCGAAGAGGTAAATCGTATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTTAAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTTAAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACAGGTT 1680
QY 1681 CGCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTACAAACGCAAGCTGATTTGCTTACCCGATGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTACAAACGCAAGCTGATTTGCTTACCCGATGCTGTT 1860
QY 1861 GAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACGTAATCTTACCTGCGCTGATTTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGTAATCTTACCTGCGCTGATTTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATCAAAAACGCAAGTACGACATTAAGAGACGTGGCGTGGATTAACCGCTGGCGGAAACGG 2100
Db 2041 TATGAAAACGCAAGTACGACATTAAGAGACGTGGCGTGGATTAACCGCTGGCGGAAACGG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAAACAAACATGAATGACAGACCGGATTC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAAACAAACATGAATGACAGACCGGATTC 2160
QY 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGTGTCACATCGACACCAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGTGTCACATCGACACCAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCATCAACCGTTCTGGTGGCTGCTGAGCGAGGTATTAACCGCCAGT 2340
Db 2281 AAGGTCACCATCAACCGTTCTGGTGGCTGCTGAGCGAGGTATTAACCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAGGTCG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGGCTAGCGCTGAAGTCTTTAGAGGAAG 2460
Db 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGGCTAGCGCTGAAGTCTTTAGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTTGCCGACCATGGAACCGCCAGAAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTTGCCGACCATGGAACCGCCAGAAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTCGCTACTGCGGTCATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTCGCTACTGCGGTCATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGTTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAAACAACCAACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAATAAACAACCAACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGTGGGCCACGCGGCCCTCGAGGCCGG 2741

Db 2701 GGATCCTCCAGTCCCGGCAAGCCCGCGCCTGGTGGGAGG 2741

```

RESULT 7
US-10-149-472-5
; Sequence 5, Application US/10149472
; Publication No. US20040029204A1
; GENERAL INFORMATION:
; APPLICANT: GOUBIN-GRAMATICA, FRANCOISE
; APPLICANT: DUCOMMUN, BERNARD
; APPLICANT: PREVOST, GREGOIRE
; TITLE OF INVENTION: METHOD FOR OBTAINING HUMAN CDC25 PHOSPHATASES AND METHOD
; TITLE OF INVENTION: FOR IDENTIFYING HUMAN CDC25 PHOSPHATASE MODULATORS
; FILE REFERENCE: bml-427.065
; CURRENT APPLICATION NUMBER: US/10/149,472
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: PCT/FR00/03496
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: FR 99/15722
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: FR 00/06883
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: FR 00/12008
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 8101
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: JM109 / pMAL - Hs Cdc25C strain
US-10-149-472-5

```

Query Match	82.0%	Score 2705.8	DB 17	Length 8101
Best Local Similarity	59.9%	Pred. No. 0		
Matches 2707	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	1	CCGACACCATCGAATGCTGCAAAACCTTTCGCGGTATGGCATGATAGGCCCGGAAGAGA	60	
Db	1	CCGACACCATCGAATGCTGCAAAACCTTTCGCGGTATGGCATGATAGGCCCGGAAGAGA	60	
QY	61	GTCAATTACGGGTGGTGAATGTGAAACCAGTAACTGTTATACGATGTGCGAGAGTATCGCG	120	
Db	61	GTCAATTACGGGTGGTGAATGTGAAACCAGTAACTGTTATACGATGTGCGAGAGTATCGCG	120	
QY	121	GTGCTCTTTATCAGACCGTTTCCGCGGTGTGAACAGGCGAGCCACGTTTCTCGAAAA	180	
Db	121	GTGCTCTTTATCAGACCGTTTCCGCGGTGTGAACAGGCGAGCCACGTTTCTCGAAAA	180	
QY	181	CGCGGAAAAAAGTGGAGCGCGGATGCGGAGCTGAAATTACATTCCCAACCGCGTGGCAC	240	
Db	181	CGCGGAAAAAAGTGGAGCGCGGATGCGGAGCTGAAATTACATTCCCAACCGCGTGGCAC	240	
QY	241	AACAACTGGCGGGCAACACAGTCGTTGCTGATTTGGCGTTGGCACTCTCAAGTCTGGCCCTGC	300	
Db	241	AACAACTGGCGGGCAACACAGTCGTTGCTGATTTGGCGTTGGCACTCTCAAGTCTGGCCCTGC	300	
QY	301	ACGCGCGCTCGCAAAATTGTCGCGCGGATTAATACTCTCGCGCGATCAACTGGGTGCCACGG	360	
Db	301	ACGCGCGCTCGCAAAATTGTCGCGCGGATTAATACTCTCGCGCGATCAACTGGGTGCCACGG	360	
QY	361	TGGTGTGTTCGATGGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC	420	
Db	361	TGGTGTGTTCGATGGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC	420	
QY	421	TTCTCCGCGAAACCGCTCAGTGGGCTGATCAATTAACATATCGCTGGATGACACAGATGCCA	480	
Db	421	TTCTCCGCGAAACCGCTCAGTGGGCTGATCAATTAACATATCGCTGGATGACACAGATGCCA	480	
QY	481	TTGCTGTGAAGCTGCGCTGCACATAATGTTCCGCGGTTATTTCTTGATGTTCTTGACCAGA	540	

481	Db	TTGCTGTGGAAAGCTGCCTGCACTAAATGTTCCGGCGGTATATTTCTTGATGTTCTCTGTGACCCAGA	541
541	Qy	CACCCATCAACAGTATTAATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
	Db		
	Db	CACCCATCAACAGTATTAATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
601	Qy	TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAAGTTCTGTCTCGG	660
601	Db	TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAAGTTCTGTCTCGG	660
661	Qy	CGCGTCTCGCTCTGGCTGGCGATCAAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
661	Db	CGCGTCTCGCTCTGGCTGGCGATCAAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
721	Qy	CGGAAACGGGAAAGGCGACTGGAGTGCCATGTCCGGTTTTTCAAACAAACCATGCAAAATGCTGA	780
721	Db	CGGAAACGGGAAAGGCGACTGGAGTGCCATGTCCGGTTTTTCAAACAAACCATGCAAAATGCTGA	780
781	Qy	ATGAGGGCATGTTTCCCACTCGGATGCTGGTTTGCCACGATCAGATGGCGCTGGGCGCAA	840
781	Db		
841	Qy	ATGAGGGCATGTTTCCCACTCGGATGCTGGTTTGCCACGATCAGATGGCGCTGGGCGCAA	840
841	Db		
841	Qy	TGCGCGCCATTACCGAGTCCGGGCTGGCGGTTGGTGCGGATATCTCGTACTGGGATACG	900
841	Db	TGCGCGCCATTACCGAGTCCGGGCTGGCGGTTGGTGCGGATATCTCGTACTGGGATACG	900
901	Qy	ACGATACGGAAGACAGTCTATGTTTATATCCCGCGTTTAAACCAATCAAAACAGATTTTC	960
901	Db	ACGATACGGAAGACAGTCTATGTTTATATCCCGCGTTTAAACCAATCAAAACAGATTTTC	960
961	Qy	GCCTGCTGGGGCAAAACAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
961	Db	GCCTGCTGGGGCAAAACAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
1021	Qy	AGGCAATACAGCTTTGCGCGTCTCACTGFTGAAAAGAAAAACACCCCTGGCGCCCAATA	1080
1021	Db	AGGCAATACAGCTTTGCGCGTCTCACTGFTGAAAAGAAAAACACCCCTGGCGCCCAATA	1080
1081	Qy	CGCAAAACCGCTCTCCCGCGGTGGCGCATTCATTAATGCGAGCTGGCAGCAGAGTTT	1140
1081	Db	CGCAAAACCGCTCTCCCGCGGTGGCGCATTCATTAATGCGAGCTGGCAGCAGAGTTT	1140
1141	Qy	CCGCACTGGAAAGCGGCGAGTGACGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
1141	Db	CCGCACTGGAAAGCGGCGAGTGACGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
1201	Qy	GCACAAATTCATGTTTTGACAGCTTATCATCGACTGACCGGTGCAACCAATGCTTCTGGCG	1260
1201	Db	GCACAAATTCATGTTTTGACAGCTTATCATCGACTGACCGGTGCAACCAATGCTTCTGGCG	1260
1261	Qy	TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGGTGCTAAATCACTGCAATAATTCG	1320
1261	Db	TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGGTGCTAAATCACTGCAATAATTCG	1320
1321	Qy	TGTCGCTCAAGGCGCACTCCCGTCTGGATAATGTTTTTTTGGCGCGCATCAATAACGGTT	1380
1321	Db	TGTCGCTCAAGGCGCACTCCCGTCTGGATAATGTTTTTTTGGCGCGCATCAATAACGGTT	1380
1381	Qy	CTGGCAAAATTTCTGAAATGAGCTGTTTGACAAATTAATCATTCGGCTCGTATAATGTGNGGA	1440
1381	Db	CTGGCAAAATTTCTGAAATGAGCTGTTTGACAAATTAATCATTCGGCTCGTATAATGTGNGGA	1440
1441	Qy	ATTGTGAGCGGATAACAAATTTCAACAGGAAAACAGCCAGTCCGTTTAGTGTGTTTTTCACGA	1500
1441	Db	ATTGTGAGCGGATAACAAATTTTCACAGGAAAACAGCCAGTCCGTTTAGTGTGTTTTTCACGA	1500
1501	Qy	GCATTTCAACCAAACAGGACCATAGATTATGAAAACCTGAAGAAGGTAACTGGTAACTCGG	1560
1501	Db	GCATTTCAACCAAACAGGACCATAGATTATGAAAACCTGAAGAAGGTAACTGGTAACTCGG	1560
1561	Qy	ATTAAACGGGATTAAGGCTATAACCGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT	1620
1561	Db	ATTAAACGGGATTAAGGCTATAACCGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT	1620

QY 1621 ACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCACACAGGTT 1680
Db |||||||
QY 1621 ACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCACACAGGTT 1680
Db |||||||
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db |||||||
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db |||||||
QY 1741 GCTCAATCTGGCCCTGTGGCTGAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db |||||||
QY 1801 CCGTTTACCTGGATGCGGTAGCTTAAACGCAAGCTGATTTGCTTACCGATCCCTGTT 1860
Db |||||||
QY 1801 CCGTTTACCTGGATGCGGTAGCTTAAACGCAAGCTGATTTGCTTACCGATCCCTGTT 1860
Db |||||||
QY 1861 GAAGGCTTATCGCTGATTTATAACAAAGATCTGCTCCGAAACCCGCAAAACCTGGGAA 1920
Db |||||||
QY 1921 GAGATCCCGGCTCGATAAAGAACTGAAGCGAAAGGTAAGAGCGCTGATGTTCAAC 1980
Db |||||||
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGTGATTTGCTGTACGGGGGTATGCGTTCAAG 2040
Db |||||||
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGTGATTTGCTGTACGGGGGTATGCGTTCAAG 2040
Db |||||||
QY 2041 TATGAAGACGCAAGTACGACATTAAGACGTGGCGGTGGATAAGCTGGCGGAAAGG 2100
Db |||||||
QY 2041 TATGAAGACGCAAGTACGACATTAAGACGTGGCGGTGGATAAGCTGGCGGAAAGG 2100
Db |||||||
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATCAATCAATCAACACCTGATAC 2160
Db |||||||
QY 2161 TCCATCGAAGAGTGCCTTTAATAAGCGCAAGACAGCGATGACCATCAACGCGCCGTGG 2220
Db |||||||
QY 2221 GCATGGTCCACATCGACACGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db |||||||
QY 2281 AAGGTCACACCATCAAAACCGTTCTGTGGCGTCTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db |||||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAACTATCTGACTGATGAAGGTCGT 2400
Db |||||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGCGGTAGCGTGAAGTCTTACGAGGAAG 2460
Db |||||||
QY 2461 TTGGGAAAGATCCACGATTTCCCGCCACCATGGAAGAACCCGAGAAAGGTGAATCATG 2520
Db |||||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
Db |||||||
QY 2581 GCCAGGGTCGTGACATGTCATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db |||||||
QY 2641 AACAAACAAACAATAAACAATAACACACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db |||||||

QY 2701 GGATCCTCT 2709
Db |||||||
QY 2701 GGATCCTCT 2709
Db |||||||
RESULT 8
US-10-343-859-8
; Sequence 8, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; PRIOR FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-343-859-8

Query Match 81.8%; Score 2701; DB 19; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGCACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGAAGAGA 60
Db |||||||
QY 1 CGCACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGAAGAGA 60
Db |||||||
QY 61 GTCATTCAGGGTGGTGAATGTGAACCCAGTAAAGTATACGATGTCGACAGTATGCG 120
Db |||||||
QY 61 GTCATTCAGGGTGGTGAATGTGAACCCAGTAAAGTATACGATGTCGACAGTATGCG 120
Db |||||||
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAAA 180
Db |||||||
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAAA 180
Db |||||||
QY 191 CGCGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTAATTCCTCCCAACCGCGTGGC 240
Db |||||||
QY 191 CGCGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTAATTCCTCCCAACCGCGTGGC 240
Db |||||||
QY 241 AACAACTCGCGGCAACAGTCGTTGCTGATTGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db |||||||
QY 241 AACAACTCGCGGCAACAGTCGTTGCTGATTGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db |||||||
QY 301 ACGGCCGCTCGCAAAATTTCTCGCGCGATTAATCTCGCGCGGATCAATCGGCGTGCAGCG 360
Db |||||||
QY 301 ACGGCCGCTCGCAAAATTTCTCGCGCGATTAATCTCGCGCGGATCAATCGGCGTGCAGCG 360
Db |||||||
QY 361 TGGTGGTTCGATGTTAGAACGAGCGCGTGAAGCCCTGAAGCGCGGTGCAATTC 420
Db |||||||
QY 361 TGGTGGTTCGATGTTAGAACGAGCGCGTGAAGCCCTGAAGCGCGGTGCAATTC 420
Db |||||||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATCTCCGCTGATGACCAAGGATGCCA 480
Db |||||||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATCTCCGCTGATGACCAAGGATGCCA 480
Db |||||||
QY 481 TTGCTGTGGAAGCTGCTGCACATTAATGTTCCGGCGTATTTCTTGTATGTTCTGACCCAGA 540
Db |||||||
QY 481 TTGCTGTGGAAGCTGCTGCACATTAATGTTCCGGCGTATTTCTTGTATGTTCTGACCCAGA 540
Db |||||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGTTACGGCACTGGGCGTGGAGCATC 600
Db |||||||

Db 541 CACCCNATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGATGGCGGTGAGCATC 600
Qy 601 TGGTCGCATTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCACTTAAGCTTCTGCTCGG 660
Db 601 TGGTCGCATTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCACTTAAGCTTCTGCTCGG 660
Qy 661 CCGGCTCGGCTGTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CCGGCTGTGGCTGTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Qy 721 CCGAAGCGGAAGCGGACGTGAGTGCCATGTCGGGTTTCAAAACCAATGCAATGCTGA 780
Db 721 CCGAAGCGGAAGCGGACGTGAGTGCCATGTCGGGTTTCAAAACCAATGCAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGCGAA 840
Qy 841 TSGCGGCCATTACCGAGTCCGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TSGCGGCCATTACCGAGTCCGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCACCATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGAGCTGACGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGAGCTGACGACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTTAGCTCACTCAATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTTAGCTCACTCAATTAG 1200
Qy 1201 GCACAATTCATGTTTGACAGCTTATCATGACGCTGACGCTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATGACGCTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGCTGTAATCACTGCATAATTCTG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGCTGTAATCACTGCATAATTCTG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAATACCGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAATACCGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Qy 1441 ATTGTGAGCGGATTAACAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTACGA 1500
Db 1441 ATTGTGAGCGGATTAACAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTACGA 1500
Qy 1501 GCACCTTCAACAGGACCATAGATTATGAAGACTGAAGAGTAACCTGGTAATCTGG 1560
Db 1501 GCACCTTCAACAGGACCATAGATTATGAAGACTGAAGAGTAACCTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680

Qy 1681 GGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACTTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGCTTACC CGATCGCTGTT 1860
Db 1801 CCGTTTACTTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGCTTACC CGATCGCTGTT 1860
Qy 1861 GAACGGTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCCGCAAAAACCTGGGAA 1920
Db 1861 GAACGGTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCCGCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACCGCAAGTACGACATTTAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACCGCAAGTACGACATTTAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAACTGCTTTTAAAGCGAAACAGCGATGACCATCAACCGCCCGTGG 2220
Db 2161 TCCATCGCAGAACTGCTTTTAAAGCGAAACAGCGATGACCATCAACCGCCCGTGG 2220
Qy 2221 GCATGCTCAAACATCGACACAGCAAAAGTAAATTTATGGTGTAAACGCTGCTGCGACCTTC 2280
Db 2221 GCATGCTCAAACATCGACACAGCAAAAGTAAATTTATGGTGTAAACGCTGCTGCGACCTTC 2280
Qy 2281 AAGGCTCAACCATTCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAACCGCCGCACT 2340
Db 2281 AAGGCTCAACCATTCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAACCGCCGCACT 2340
Qy 2341 CCGAAACAAAGAGCTGCAAAAGAGTTTCTCGAAAACCTATCTGCTGACTGATGAGGCTCTG 2400
Db 2341 CCGAAACAAAGAGCTGCAAAAGAGTTTCTCGAAAACCTATCTGCTGACTGATGAGGCTCTG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGGTGCGGTAGCGCTGAAAGCTTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGGTGCGGTAGCGCTGAAAGCTTTTACGAGGAAGAG 2460
Qy 2461 TTTGGCAAGAGTCCAGTATTTGCGGCCACCATGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTTGGCAAGAGTCCAGTATTTGCGGCCACCATGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGATGCGTGGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGATGCGTGGCTGCTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACGCTGCTGATGAAGCCCTGAAAGACCGCGCAGACTAATTTGAGGCTCG 2640
Db 2581 GCCAGCGTCTGACGCTGCTGATGAAGCCCTGAAAGACCGCGCAGACTAATTTGAGGCTCG 2640
Qy 2641 AACACAACAACAATAACAAATTAACAACTCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACACAACAACAATAACAAATTAACAACTCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCTCT 2709
Db 2701 GGATCTCTCT 2709

RESULT 9

US-10-343-859-9
; Sequence 9, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; PRIOR FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9191
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-muts
; OTHER INFORMATION: plasmid
US-10-343-859-9

Query Match 81.8%; Score 2698; DB 19; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGCAGAAACCTTTCGGCGGTATGCGCATGATAGGCCCGCGAAGAGA	60
DB	1	CCGACACCATCGAATGGTGCAGAAACCTTTCGGCGGTATGCGCATGATAGGCCCGCGAAGAGA	60
QY	61	GTCATTCAGGCTGGTGAATGTGAACACAGTAACTTATACCATGTCGCGAGTATGCG	120
DB	61	GTCATTCAGGCTGGTGAATGTGAACACAGTAACTTATACCATGTCGCGAGTATGCG	120
QY	121	GTGCTCTTATCAGACCGGTTTCGGCGGTGTTGAAACAGCCAGCCACGCTTCTCGGAGAAA	180
DB	121	GTGCTCTTATCAGACCGGTTTCGGCGGTGTTGAAACAGCCAGCCACGCTTCTCGGAGAAA	180
QY	181	CCGGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTTACATTTCCCAACCGCTGGCAC	240
DB	181	CCGGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTTACATTTCCCAACCGCTGGCAC	240
QY	241	ACAACCTGCGGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	ACAACCTGCGGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTGCAGAAATTTGCGCGGATTAATCTCGCGCGATCAATCGGCTGCCAGCG	360
DB	301	ACGCGCGTGCAGAAATTTGCGCGGATTAATCTCGCGCGATCAATCGGCTGCCAGCG	360
QY	361	TGGTGTGTCGATGTGATAGCAAGCGCGTCCAGCCCTGTAAAGCGCGGTGCACAATC	420
DB	361	TGGTGTGTCGATGTGATAGCAAGCGCGTCCAGCCCTGTAAAGCGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGGCTGGATGACAGGATGCCA	480
DB	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGGCTGGATGACAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCTCAGTAAATTTTCGGCGTTATTTCTGATGCTCTGACCCAGA	540
DB	481	TTGCTGTGGAAGCTGCTCAGTAAATTTTCGGCGTTATTTCTGATGCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACCGACTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACCGACTGGCGGTGGAGCATC	600
QY	601	TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCATTTAAATTTCTGCTCGG	660
DB			

DB	601	TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCATTTAAATTTCTGCTCGG	660
QY	661	CGCGTCTCGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTCGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGCGCATGCTTCCCACTCGATGCTGGTTGCGCAACGATCAGATGCGCTGGCGCGAA	840
DB	781	ATGAGGCGCATGCTTCCCACTCGATGCTGGTTGCGCAACGATCAGATGCGCTGGCGCGAA	840
QY	841	TGCGCGCATTAACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGCGCATTAACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTCAACTCTCTCAGGCGCAGGCGTGA	1020
DB	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTCAACTCTCTCAGGCGCAGGCGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGCGAGTGGCAGCAGAGTTT	1140
DB	1081	CGCAACCGCTCTCTCCCGCGGTTGGCGGATTAATTAATGCGAGTGGCAGCAGAGTTT	1140
QY	1141	CCCGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTAGTTAGCTCACTCATTAG	1200
DB	1141	CCCGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTAGTTAGCTCACTCATTAG	1200
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTCTGGCG	1260
DB	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTCTGGCG	1260
QY	1261	TCAGCAGCCATCGAAGCTGTGATGCGTGTGAGGTGTAATCATCTGATCAATTTTCG	1320
DB	1261	TCAGCAGCCATCGAAGCTGTGATGCGTGTGAGGTGTAATCATCTGATCAATTTTCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCGTTCTGATTAATGTTTTCGCGCGACATCAATACGTT	1380
DB	1321	TGTCGCTCAAGCGCACCTCCGTTCTGATTAATGTTTTCGCGCGACATCAATACGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500
DB	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAACTGGAAGAGGTAAATCTGGTAAATCTGG	1560
DB	1501	GCACCTTCAACCAAGGACCATAGATTATGAACTGGAAGAGGTAAATCTGGTAAATCTGG	1560
QY	1561	ATTAAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT	1620
DB	1561	ATTAAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCGCCAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCGCCAGGTT	1680
QY	1681	GGGGCAATCGGCGTGGCGCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC	1740
DB	1681	GGGGCAATCGGCGTGGCGCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC	1740


```
QY 1741 GCTCAATCTGGCTGTGTGGCTGAATCACTCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAATCACTCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTAGCTTTACACGGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCGGTAGCTTTACACGGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAAAGCGTTATCGCTGATTTATAACAAAGATCTGTCGCGAACCCGCCAAACCTCGGAA 1920
Db 1861 GAAAGCGTTATCGCTGATTTATAACAAAGATCTGTCGCGAACCCGCCAAACCTCGGAA 1920
QY 1921 GAGATCCCGGCTCGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGGCTCGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGACCGTACTTACCTGGCGGTGATTTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGACCGTACTTACCTGGCGGTGATTTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGGTGATTAAGCGTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGGTGATTAAGCGTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTAAATAAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTAAATAAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGTTGTAAGCGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGTTGTAAGCGTACTGCCGACCTTC 2280
QY 2281 AAGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGGAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGGAGGTATTAACCGCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATCTGCTGACTGATGAAGGTTCTG 2400
QY 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAAG 2460
Db 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAAG 2460
QY 2461 TTGGGAAAGATCCACGTTATTTGCCGCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTTATTTGCCGCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGCGTGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGCGTGTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
Db 2581 GCCAGCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
QY 2641 AACAAACAAACAAATACAAATACAAACAACTCGGATCGAGGAAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATACAAATACAAACAACTCGGATCGAGGAAAGGATTCAGAAATTC 2700
QY 2701 GGATCC 2706
Db 2701 GGATCC 2706
```

RESULT 10

US-10-263-153-40

; Sequence 40, Application US/10263153

```
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 7112
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3189)
; OTHER INFORMATION: pMBP-c2X-Toxop30del11 (52-214aa)
; US-10-263-153-40
```

```
Query Match 81.5%; Score 2690.4; DB 18; Length 7112;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 CCGACACCATCGAATGTGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGTGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGAAGAGA 60
QY 61 GTCAATTCAGGTGGTGAATGTGAACCAAGTAACTGATACGATGTGCGAGATGATGCGG 120
Db 61 GTCAATTCAGGTGGTGAATGTGAACCAAGTAACTGATACGATGTGCGAGATGATGCGG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTGTAACCAAGCGCCAGCCAGCTTTCTGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTGTAACCAAGCGCCAGCCAGCTTTCTGCGAAA 180
QY 181 CCGGGGAAAAGTGAAGCGCGCATGCGGAGCTGAAATTAATTCCTCCAAACCGGTGGCAC 240
Db 181 CCGGGGAAAAGTGAAGCGCGCATGCGGAGCTGAAATTAATTCCTCCAAACCGGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGCTGATTGGCGGTGGCACCTCCAGTCTGCGCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGCTGATTGGCGGTGGCACCTCCAGTCTGCGCCTGC 300
QY 301 ACGGCGCTCGCAAAATGTGCGCGGATTAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGGCGCTCGCAAAATGTGCGCGGATTAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGTGTGTGTCGATGTTAGAACGAAGCGCGTGAAGCGCTGTAAGCGGCGGTGCACAATC 420
Db 361 TGTGTGTGTCGATGTTAGAACGAAGCGCGTGAAGCGCTGTAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTAATCTCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTAATCTCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTAGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTAGCGACTGGGCGTGGAGCATC 600
QY 601 TGTGTCGATTTGGTCAACAGCAAAATCGCTGTTAGCGGGCCCAATTAAGTTCTGCTCGG 660
Db 601 TGTGTCGATTTGGTCAACAGCAAAATCGCTGTTAGCGGGCCCAATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
```

Db 661 |||||CGCGTCTGCGTGGCTGGAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTTGCCAACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTTGCCAACGATCAGATGGCGTGGCGCAA 840
Qy 841 TGCGGCAATTACAGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTGTGGGATACG 900
Db 841 TGCGGCAATTACAGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACACCATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACACAGCGTGGACCGGTTGCTGCAATCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAACACAGCGTGGACCGGTTGCTGCAATCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAACACACCTGGGCGCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAACACACCTGGGCGCAATA 1080
Qy 1081 CGCAAAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGGCACACAGGTTT 1140
Db 1081 CGCAAAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGGCACACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAGGTCGACCAATGCTTCTGGG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAGGTCGACCAATGCTTCTGGG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGCTTCGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGCTTCGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACATTAATCATCGGCTCGTATATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACATTAATCATCGGCTCGTATATGTGTGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTTCACAGAAACAGCCAGTCGGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCACAGAAACAGCCAGTCGGTTTAGTGTTCACGA 1500
Qy 1501 GCACCTTCAACAAAGGCCATAGATTATGAAAACCTGAAGAGGTAACCTGGTAATCTGG 1560
Db 1501 GCACCTTCAACAAAGGCCATAGATTATGAAAACCTGAAGAGGTAACCTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATACCGGTCGCTGAAATCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATACCGGTCGCTGAAATCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAGAGAAATTCGCCAGGTT 1680
Qy 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTCGGGCACACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTCGGGCACACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
|||||

Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTTACTCTGGATGCGGTACGTTTCAACCGGCAAGCTGATTGTTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTTACTCTGGATGCGGTACGTTTCAACCGGCAAGCTGATTGTTTACCCGATCGCTGTT 1860
Qy 1861 GAAGGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAACGAAAGTAAAGAGCGGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAACGAAAGTAAAGAGCGGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGGCCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAGAACGCAAGTAGACATTAAGACCTGGGCTGGATTAACGCTGGGCGGAAAGCG 2100
Db 2041 TATGAAGAACGCAAGTAGACATTAAGACCTGGGCTGGATTAACGCTGGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGCTGATTAATAAACAACAACATGAAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGCTGATTAATAAACAACAACATGAAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCCCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACACCATCCAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACCGCCCACT 2340
Db 2281 AAGGTCACACCATCCAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACCGCCCACT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGCTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGCTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGGAAAGATCCACGATTTGCGCCCACTATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGATTTGCGCCCACTATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGCTGCTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGGTCTGACGTGCGATGAAGCCTTGAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCTGACGTGCGATGAAGCCTTGAAGACGCGCAGACTTAATTCGAGCTCG 2640
Qy 2641 AACCAACAAACAAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700

RESULT 11
US-10-263-153-30
; Sequence 30, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.

; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 7259
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3336)
; OTHER INFORMATION: pMBP-c2X-Toxop30del14del18 (83-294aa)
US-10-263-153-30

Query Match 81.5%; Score 2690.4; DB 18; Length 7259;
Best Local Similarity 99.8%; Fred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60

QY 61 GTCAATTCAGGGTGTGAATGTGAACCAAGTAACGTTTATACGATGTCCAGAGTATGCCG 120
DB 61 GTCAATTCAGGGTGTGAATGTGAACCAAGTAACGTTTATACGATGTCCAGAGTATGCCG 120

QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCGAGCCAGCCACGTTTCTCGGAAA 180
DB 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCGAGCCAGCCACGTTTCTCGGAAA 180

QY 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240

QY 241 AACAACTGGCGGCAACAGTCGTGTGTGATGGCGTTGCCACTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAACAGTCGTGTGTGATGGCGTTGCCACTCCAGTCTGGCCCTGC 300

QY 301 ACGCGCGTCCGCAATTTGTCCGCGGATTAATCTCGCGCGATCAACTGGTGGCCAGCG 360
DB 301 ACGCGCGTCCGCAATTTGTCCGCGGATTAATCTCGCGCGATCAACTGGTGGCCAGCG 360

QY 361 TGGTGTGTGATGATGATAGAACGAGCGCGTGAAGCTGTGAAGCGCGGTGCACAATC 420
DB 361 TGGTGTGTGATGATGATAGAACGAGCGCGTGAAGCTGTGAAGCGCGGTGCACAATC 420

QY 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCGCTGGATGACGAGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCGCTGGATGACGAGATGCCA 480

QY 481 TTGCTGTGGAAGCTCCGTCACATAATGTTCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGGAAGCTCCGTCACATAATGTTCGGCGTTATTTCTTGATGTCTCTGACCAGA 540

QY 541 CACCCATCAACAGTATATTTTCCCATGAAGACGCTACCGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTTCCCATGAAGACGCTACCGACTGGCGGTGGAGCATC 600

QY 601 TGGTTCGCAATTTGGGTACACAGCAATCCGCTGTAGCGGCGCCATTAAGTCTGTCTCGG 660
DB 601 TGGTTCGCAATTTGGGTACACAGCAATCCGCTGTAGCGGCGCCATTAAGTCTGTCTCGG 660

QY 661 CGGCTTCGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGGCTTCGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAACGGGAAGCGACTGAGTGCCTATCTCCGTTTTTCAACAAACCATGCAATGCTGA 780
DB 721 CGGAACGGGAAGCGACTGAGTGCCTATCTCCGTTTTTCAACAAACCATGCAATGCTGA 780

QY 781 ATGAGGCATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGCGCTGGCGCGAA 840
DB 781 ATGAGGCATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGCGCTGGCGCGAA 840

QY 841 TGCAGCCATTAACGAGTCCGGGTCCGGGTCCGGGTCCGGGTCCGGGTCCGGGTCCGGGTACG 900
DB 841 TGCAGCCATTAACGAGTCCGGGTCCGGGTCCGGGTCCGGGTCCGGGTCCGGGTACG 900

QY 901 ACGATACGGAAGACAGCTCATGTTATATCCGCCGTTAAACCAATCAACAGAGATTTTC 960
DB 901 ACGATACGGAAGACAGCTCATGTTATATCCGCCGTTAAACCAATCAACAGAGATTTTC 960

QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020

QY 1021 AGGCAATACGCTGTTGCCGCTCTCACTGGTGAAGAAAAAACCCACCTCGCGCCCAATA 1080
DB 1021 AGGCAATACGCTGTTGCCGCTCTCACTGGTGAAGAAAAAACCCACCTCGCGCCCAATA 1080

QY 1081 CGCAAAACCGCTCTCCCGCGCGTGGCCGATTAATTAATGAGCTGGACGACGAGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCGTGGCCGATTAATTAATGAGCTGGACGACGAGTTT 1140

QY 1141 CCGACTTGAAGCGGCGAGTGAACGCAACGCAATTAATGTAAGTTAGCTCACTATTAG 1200
DB 1141 CCGACTTGAAGCGGCGAGTGAACGCAACGCAATTAATGTAAGTTAGCTCACTATTAG 1200

QY 1201 GCAAAATTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCAAAATTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260

QY 1261 TCAGGACGACCATCGAAGCTGTGATGCTGAGTAAATCACTGCAATTAATTCG 1320
DB 1261 TCAGGACGACCATCGAAGCTGTGATGCTGAGTAAATCACTGCAATTAATTCG 1320

QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAAAACGTT 1380
DB 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAAAACGTT 1380

QY 1381 CTGGCAATATTTGAAATGAGCTGTGCAATTAATCACTGCGCTCGTAAATGTGTGGA 1440
DB 1381 CTGGCAATATTTGAAATGAGCTGTGCAATTAATCACTGCGCTCGTAAATGTGTGGA 1440

QY 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTCCGTTTGGTTTTCACGA 1500
DB 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTCCGTTTGGTTTTCACGA 1500

QY 1501 GCATTTCAACCAAGGACCATAGCATATGAAACTGAAAGAGGTAAACTGGTAAATCTCG 1560
DB 1501 GCATTTCAACCAAGGACCATAGCATATGAAACTGAAAGAGGTAAACTGGTAAATCTCG 1560

QY 1561 ATTAAACGGGATAAAGGCTATACCGTCTCGCTGAAGTAAAGAAATTCGAAAGAT 1620
DB 1561 ATTAAACGGGATAAAGGCTATACCGTCTCGCTGAAGTAAAGAAATTCGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCAACCGTTCAGCATCCGATTAACCTGAAAGAGAAATTCACACAGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTCAGCATCCGATTAACCTGAAAGAGAAATTCACACAGTT 1680

QY 1681 GGGGCAACTGGCGATGGCCCTGACATTAATCTCTGCGGCAACGCGCTTGGTGGCTAC 1740
DB 1681 GGGGCAACTGGCGATGGCCCTGACATTAATCTCTGCGGCAACGCGCTTGGTGGCTAC 1740

QY 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGCAAGCGCTTCAGGACAGCTGAT 1800
DB 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGCAAGCGCTTCAGGACAGCTGAT 1800

QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGGAAGCTGATTGCTTACCGGATCGCTGT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGGAAGCTGATTGCTTACCGGATCGCTGT 1860

QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920

[illegible]

```

RESULT 12
US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0

```

```
; SEQ ID NO 35
;
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3399)
; OTHER INFORMATION: pmBP-c2X-Toxop3odel10 (52-284aa)
US-10-263-153-35
```

Query Match	81.5%	Score 2690.4	DB 18	Length 7322
Best Local Similarity	99.8%	Pred. No. 0		
Matches 2694	Conservative 0	Mismatches 6	Indels 0	Gaps 0
Qy 1	CCGACACATCGAATGTCGAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAAGAGA	60		
Db 1	CCGACACATCGAATGTCGAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAAGAGA	60		
Qy 61	GTCAATTCAGGCTGGTGAATGTGAACACGATAAGTTATACGATGTCGCAGAGTATCCG	120		
Db 61	GTCAATTCAGGCTGGTGAATGTGAACACGATAAGTTATACGATGTCGCAGAGTATCCG	120		
Qy 121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACAGGCGAGCCAGCTTCTGCGAATA	180		
Db 121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACAGGCGAGCCAGCTTCTGCGAATA	180		
Qy 181	CGCGGGAAGAAGTGAAGCGCGATGCGGAGCTGAATTACATTCACACCGCTGGCAC	240		
Db 181	CGCGGGAAGAAGTGAAGCGCGATGCGGAGCTGAATTACATTCACACCGCTGGCAC	240		
Qy 241	AACAACTGGCGGGCAACAGTCGTTGCTGATGTGCGTGGCCACTTCAGTCTGGCCCTGC	300		
Db 241	AACAACTGGCGGGCAACAGTCGTTGCTGATGTGCGTGGCCACTTCAGTCTGGCCCTGC	300		
Qy 301	ACGCGCGTCGCAAAATGTGCGCGCATTAATCTCGCGCGATCACTGGTGGCCAGCG	360		
Db 301	ACGCGCGTCGCAAAATGTGCGCGCATTAATCTCGCGCGATCACTGGTGGCCAGCG	360		
Qy 361	TGCTGGTGTGATGCTAGAACGAGCGGCTGGAAGCTGTAAAGCGCGGTGCACAATC	420		
Db 361	TGCTGGTGTGATGCTAGAACGAGCGGCTGGAAGCTGTAAAGCGCGGTGCACAATC	420		
Qy 421	TTCTCGCGCAACGCGTCAGTGGGCTGATTAATACTATCCGCTGGATGACCCAGGATGCCA	480		
Db 421	TTCTCGCGCAACGCGTCAGTGGGCTGATTAATACTATCCGCTGGATGACCCAGGATGCCA	480		
Qy 481	TTGCTGTGAAGCTGCGCTGCACATAATGTTCCGGCGTTATTTCCTGATGTCCTGACACAGA	540		
Db 481	TTGCTGTGAAGCTGCGCTGCACATAATGTTCCGGCGTTATTTCCTGATGTCCTGACACAGA	540		
Qy 541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGCTAGCGACTGGCGGTGAGGATC	600		
Db 541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGCTAGCGACTGGCGGTGAGGATC	600		
Qy 601	TGCTCGCATTTGGGTCAACAGCAAAATCGCTGTTAGCGGGCCCAATTAACTCTCTCGG	660		
Db 601	TGCTCGCATTTGGGTCAACAGCAAAATCGCTGTTAGCGGGCCCAATTAACTCTCTCGG	660		
Qy 661	CGCGTCTGCGTCTGGCTGGCGATTAATAATCTCACTCGCAATCAAAATTCAGCGGATAG	720		
Db 661	CGCGTCTGCGTCTGGCTGGCGATTAATAATCTCACTCGCAATCAAAATTCAGCGGATAG	720		
Qy 721	CGGAACGGGAAGCGACTGGAGTCCCATGTCGGGTTTTCAACAACCAATGCAATGCTGA	780		
Db 721	CGGAACGGGAAGCGACTGGAGTCCCATGTCGGGTTTTCAACAACCAATGCAATGCTGA	780		
Qy 781	ATGAGGGCATCGTTCCCACTGCGATGCTGGTGTGCAACAGATCAGATGCGCTGGGCGCAA	840		
Db 781	ATGAGGGCATCGTTCCCACTGCGATGCTGGTGTGCAACAGATCAGATGCGCTGGGCGCAA	840		
Qy 841	TGCGGCCATTACCGAGTTCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATACG	900		
Db 841	TGCGGCCATTACCGAGTTCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATACG	900		

; OTHER INFORMATION: pMBP-c2X-ToxoP30del14C (52-294aa)
US-10-263-153-25

Query Match	81.5%	Score 2690.4	DB 18	Length 7352
Best Local Similarity	99.8%	Pred. No. 0		
Matches 2694	Conservative 0	Mismatches 6	Indels 0	Gaps 0
Qy	1	CCGACACCATCGAATGGTGTGAAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA	60	
Db	1	CCGACACCATCGAATGGTGTGAAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA	60	
Qy	61	GTCAATTACAGGTGGTGAATGTGAACCAAGTAACGTTTATACGATGTGCGACAGTATGCCG	120	
Db	61	GTCAATTACAGGTGGTGAATGTGAACCAAGTAACGTTTATACGATGTGCGACAGTATGCCG	120	
Qy	121	GTGTCTCTTATCAGACCGCTTTCCCGCGTGTGAACCAAGCCAGCCAGCCAGTTCCTCGGAAA	180	
Db	121	GTGTCTCTTATCAGACCGCTTTCCCGCGTGTGAACCAAGCCAGCCAGCCAGTTCCTCGGAAA	180	
Qy	181	CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAAATTAATTCACCAACCGCTGGCAC	240	
Db	181	CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAAATTAATTCACCAACCGCTGGCAC	240	
Qy	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACTTCAGTCTGGCCCTGC	300	
Db	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACTTCAGTCTGGCCCTGC	300	
Qy	301	ACGGCCGTCGCAAAATTTGTCCGCGCGATTAATCTCTCGGCGCATCACTGGGTGCGACG	360	
Db	301	ACGGCCGTCGCAAAATTTGTCCGCGCGATTAATCTCTCGGCGCATCACTGGGTGCGACG	360	
Qy	361	TGGTGGTGTGATCGTATAGAACGAAGCGCGTCGAAAGCTGTAAAGCGCGGTGCACAATC	420	
Db	361	TGGTGGTGTGATCGTATAGAACGAAGCGCGTCGAAAGCTGTAAAGCGCGGTGCACAATC	420	
Qy	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACGAGTGCCA	480	
Db	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACGAGTGCCA	480	
Qy	481	TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTTATTTCTTGATGTCTCTGACAGA	540	
Db	481	TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTTATTTCTTGATGTCTCTGACAGA	540	
Qy	541	CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGCGGTGAGGATC	600	
Db	541	CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGCGGTGAGGATC	600	
Qy	601	TGTFPGCATTTGGGTACACAGCAAAATCGCGTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG	660	
Db	601	TGTFPGCATTTGGGTACACAGCAAAATCGCGTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG	660	
Qy	661	CGCGTCTCGCTCTGGCTGGCGGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
Db	661	CGCGTCTCGCTCTGGCTGGCGGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
Qy	721	CGGAAACGGGAAGGCGACTGGAGTGCCATGTCGGGTTTTTCAACAAACCATGCAATGCTGA	780	
Db	721	CGGAAACGGGAAGGCGACTGGAGTGCCATGTCGGGTTTTTCAACAAACCATGCAATGCTGA	780	
Qy	781	ATGAGGCGATCGTTTCCCACTGCGATGTGTTGTCACACGATCAGATGCGCTGGGCGCAA	840	
Db	781	ATGAGGCGATCGTTTCCCACTGCGATGTGTTGTCACACGATCAGATGCGCTGGGCGCAA	840	
Qy	841	TGCGCGCATTTACCGAGTCGCGGCTCGCGTTGTTGCGGATATCTCGTAGTGGGATACG	900	
Db	841	TGCGCGCATTTACCGAGTCGCGGCTCGCGTTGTTGCGGATATCTCGTAGTGGGATACG	900	
Qy	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGCTTAAACCAACCATCAACAGGATTTTC	960	
Db	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGCTTAAACCAACCATCAACAGGATTTTC	960	
Qy	961	GCCTGCTGGGCGAAACCAAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020	

961	GCCTGCTGGGGCAAA	CCAGCGTGGAC	CGCGTTGCTGC	AACTCTCTCTCA	AGGCGCAC	AGGCGGTGA	1021
Qy	1021	AGGCAATCAGCTGTTG	CCCGCTCTCACT	GGTGA	AAAAACAC	CCCTGGCGCCCAATA	1080
	1021	AGGCAATCAGCTGTTG	CCCGCTCTCACT	GGTGA	AAAAACAC	CCCTGGCGCCCAATA	1080
	1081	CGMAACCGGCTCT	CCCCGGCGTTGG	CGATTCATTTA	TGACGTGG	CACGACGAGT	1140
Qy	1081	CGMAACCGGCTCT	CCCCGGCGTTGG	CGATTCATTTA	TGACGTGG	CACGACGAGT	1140
	1141	CCGACGTGGA	AAGCGGCGAGT	GACGCAACG	CAATTAATG	TGAGTTAGCTCACTCAT	1200
	1141	CCGACGTGGA	AAGCGGCGAGT	GACGCAACG	CAATTAATG	TGAGTTAGCTCACTCAT	1200
Qy	1201	GCACAA	TTCTCATGTTTG	ACAGCTTATCAT	PCGACTG	CACGCGTCA	1260
	1201	GCACAA	TTCTCATGTTTG	ACAGCTTATCAT	PCGACTG	CACGCGTCA	1260
	1261	TCAGGAC	CCCATCGGAAGCTGT	GGTATGGCTGTG	CAGGTCGTA	AAATCAGTCGAT	1320
Qy	1261	TCAGGAC	CCCATCGGAAGCTGT	GGTATGGCTGTG	CAGGTCGTA	AAATCAGTCGAT	1320
	1321	TGTCGCTCA	AGGCGCACTCC	CGTTCTTGGA	TAATGTTT	TGCGCGCAG	1380
	1321	TGTCGCTCA	AGGCGCACTCC	CGTTCTTGGA	TAATGTTT	TGCGCGCAG	1380
Qy	1381	CTGGCA	AAATATCTG	AAATGAGCTGTT	TGACAA	TTAATCAT	1440
	1381	CTGGCA	AAATATCTG	AAATGAGCTGTT	TGACAA	TTAATCAT	1440
	1441	ATTGTG	AGCGGATAACAA	TTTCA	CACAGGAAA	CAGCGAGT	1500
Qy	1441	ATTGTG	AGCGGATAACAA	TTTCA	CACAGGAAA	CAGCGAGT	1500
	1501	GCATTT	CACCAACAG	GACCATAGATAT	TGAAA	AACTG	1560
	1501	GCATTT	CACCAACAG	GACCATAGATAT	TGAAA	AACTG	1560
Qy	1561	ATTAA	CGGCGATAA	AGGCTATAAC	CGGTCTCGCT	GAACTCGGTA	1620
	1561	ATTAA	CGGCGATAA	AGGCTATAAC	CGGTCTCGCT	GAACTCGGTA	1620
	1621	ACCGGA	ATTAAGTCA	CCGTTGAGCTAC	CGGATAA	ACTGGA	1680
Qy	1621	ACCGGA	ATTAAGTCA	CCGTTGAGCTAC	CGGATAA	ACTGGA	1680
	1681	CGCGCA	ACTGGCGAT	GGCCCTGACAT	TATCTTCT	GGGCACACAC	1740
	1681	CGCGCA	ACTGGCGAT	GGCCCTGACAT	TATCTTCT	GGGCACACAC	1740
Qy	1741	GCTCA	ATCTGGCCTGT	TGGCTGAA	AACTCAC	CCCGCA	1800
	1741	GCTCA	ATCTGGCCTGT	TGGCTGAA	AACTCAC	CCCGCA	1800
	1801	CCGTTA	ACCTGGGATG	CGTACGTTT	TAACCGGCA	AGCTGAT	1860
Qy	1801	CCGTTA	ACCTGGGATG	CGTACGTTT	TAACCGGCA	AGCTGAT	1860
	1861	GAAGCG	TTATCGCTGATTT	ATAACAA	AGATCTGCT	CGCGAAC	1920
	1861	GAAGCG	TTATCGCTGATTT	ATAACAA	AGATCTGCT	CGCGAAC	1920
Qy	1921	GAGAT	CCCGCGCTGG	ATAAGAACTG	AAAGCG	AAAGGTAA	1980
	1921	GAGAT	CCCGCGCTGG	ATAAGAACTG	AAAGCG	AAAGGTAA	1980
	1981	CTGCA	GAAACCGTACTT	CACCTGG	CGCTGATTTG	CTGTCG	2040
Qy	1981	CTGCA	GAAACCGTACTT	CACCTGG	CGCTGATTTG	CTGTCG	2040
	2041	TATGAA	AACCGCAAGTAC	GA	CA	ATAA	2100
	2041	TATGAA	AACCGCAAGTAC	GA	CA	ATAA	2100


```
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGCTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGCTGG 2220
QY 2221 GCATGGTCCACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTCGGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAACACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGGCCCAAGT 2340
Db 2281 AAGGGTCAACCATCAACACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGGCCCAAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTGGTGAAGCTGTAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTGGTGAAGCTGTAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTTGGCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTTGGCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGAGTCCGCTTCTGGTATGCGGTGCTGAGTCCGCTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGAGTCCGCTTCTGGTATGCGGTGCTGAGTCCGCTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGAGCTGTGATGAAGCCCTGAAAGCGGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGAGCTGTGATGAAGCCCTGAAAGCGGCGAGACTAATTCGAGCTCG 2640
QY 2641 AACACAAACAATATCAATATCAACAACTCGGATCGAGGAAGATTTTCAGAAATC 2700
Db 2641 AACACAAACAATATCAATATCAACAACTCGGATCGAGGAAGATTTTCAGAAATC 2700
```

RESULT 14

```
US-10-263-153-20
; Sequence 20, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN.
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxo30del13C (52-300aa)
US-10-263-153-20
```

```
Query Match 81.58; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.84; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 CCGACACCATCGAATGCTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGCTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCAATTCAGGGTGTCAATGTGAACCAAGTAAAGTGTATACGATGTCGCGAGAGTATCGCG 120
Db 61 GTCAATTCAGGGTGTCAATGTGAACCAAGTAAAGTGTATACGATGTCGCGAGAGTATCGCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCGCGTGTGTGAACCAAGGCGAGCCAGCTTTCTCGCAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCGCGTGTGTGAACCAAGGCGAGCCAGCTTTCTCGCAAAA 180
QY 181 CCGGGGAAAAGTGAAGCGCGCATGCGCGAGCTGAATTAATTCACCAACCGGTGGCAC 240
Db 181 CCGGGGAAAAGTGAAGCGCGCATGCGCGAGCTGAATTAATTCACCAACCGGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTTGCCCTGCG 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTTGCCCTGCG 300
QY 301 ACGGCGCGTGCAGAAATTTGTCGCGCGAATTAATTTCTCGCGCGGATCAACTGGGTGCCACG 360
Db 301 ACGGCGCGTGCAGAAATTTGTCGCGCGAATTAATTTCTCGCGCGGATCAACTGGGTGCCACG 360
QY 361 TGGTGGTGTGATGTAGAACGAAGCGGCTGCAAGCCTGTAAGCGGCGGTGCACAAATC 420
Db 361 TGGTGGTGTGATGTAGAACGAAGCGGCTGCAAGCCTGTAAGCGGCGGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAAATTTCCGCGGTTATTTCTTGATGTCTTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAAATTTCCGCGGTTATTTCTTGATGTCTTGACACAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTAGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTAGCGACTGGGCGTGGAGCATC 600
QY 601 TGTTCGATTTGGTCAACGAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
Db 601 TGTTCGATTTGGTCAACGAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
QY 661 CGGCTCTGCTGCTGGCTGSCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Db 661 CGGCTCTGCTGCTGGCTGSCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
QY 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTTCACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCACTGCGATGCTGTTGCCAAGCATGAGTGGCGTGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCACTGCGATGCTGTTGCCAAGCATGAGTGGCGTGGCGCAA 840
QY 841 TGGCGCCATTACCGAGTCCGCGTGGTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGCCATTACCGAGTCCGCGTGGTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAAGGATTTTC 960
Db 901 ACATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAAGGATTTTC 960
QY 961 GCTGCTGGGCAAAACAGCGTGGAGCTGCTGCACTCTCTCAGGGCCAGGGCGTGA 1020
Db 961 GCTGCTGGGCAAAACAGCGTGGAGCTGCTGCACTCTCTCAGGGCCAGGGCGTGA 1020
QY 1021 AGGGCAATCAGTGTGTGCGCGTCTCACTGGTGAAGAAACCAACCTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGTGTGTGCGCGTCTCACTGGTGAAGAAACCAACCTTGGCGCCCAATA 1080
QY 1081 CGCAACACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGCTGGCACACAGGTTT 1140
```


[illegible]

Db	2161	TCATCGCAGAGCTGCTTTTAATAAAGGCGAAACAGCGATGACCAATCAACAGCCCGGTGG	2222
Qy	2221	GCATGGTCCAAACATCGACACAGCAAGAGTGAATTTATGGTGTAAACGGTACTTGCACGACCTTC	2280
Db	2221	GCATGGTCCAAACATCGACACAGCAAGAGTGAATTTATGGTGTAAACGGTACTTGCACGACCTTC	2280
Qy	2281	AAGGGTCAACCATCCAAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Db	2281	AAGGGTCAACCATCCAAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Qy	2341	CCGAACAAAGAGCTGCGAAAGAGTTCCCTCGAATACTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAACAAAGAGCTGCGAAAGAGTTCCCTCGAATACTATCTGCTGACTGATGAAGTCTG	2400
Qy	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGGAAGAG	2460
Qy	2461	TTGCGGAAAGATCCACGTATTTGCGGCCACCATCGAAACCGCCAGAAAGGTGAAATCATG	2520
Db	2461	TTGCGGAAAGATCCACGTATTTGCGGCCACCATCGAAACCGCCAGAAAGGTGAAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGCCGCTTTCTGATGTCGGTGGGTACTCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGCCGCTTTCTGATGTCGGTGGGTACTCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGTGGTCCAGACTGTCGATGAAGCCCTCGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGTGGTCCAGACTGTCGATGAAGCCCTCGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAACAACAACAAATTAACATTAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2641	AACAACAACAACAAATTAACATTAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
RESULT 15			
US-10-263-153-61			
; Sequence 61, Application US/10263153			
; Publication No. US20040067239A1			
; GENERAL INFORMATION:			
; APPLICANT: Abbott Laboratories			
; APPLICANT: Maine, Gregory T.			
; APPLICANT: Patel, Chandu B.			
; APPLICANT: Ginsburg, Sanford R.			
; APPLICANT: Bliese, Timothy R.			
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,			
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF			
; FILE REFERENCE: 6984.US.O1			
; CURRENT APPLICATION NUMBER: US/10/263,153			
; CURRENT FILING DATE: 2002-10-22			
; NUMBER OF SEQ ID NOS: 74			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 61			
; LENGTH: 7370			
; TYPE: DNA			
; ORGANISM: Toxoplasma gondii			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1528)...(3447)			
; OTHER INFORMATION: pMBP-C2X-ToxOP30MIX1			
US-10-263-153-61			

	Query Match	81.5%	Score 2690.4	DB 18	Length 7370
	Best Local Similarity	99.8%	Pred. No. 0		
	Matches 2694	Conservative 0	Mismatches 6	Indels 0	Gaps 0
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCGCGGTATGGCATGTACGCGCCCGAGAGAGA	60		
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCGCGGTATGGCATGTACGCGCCCGAGAGAGA	60		
Qy	61	GTCAATTCCAGGTGGTGAATGTGAACACCACTAACGTTATACGATGTCCGAGAGTATGCCG	120		
Db	61	GTCAATTCCAGGTGGTGAATGTGAACACCACTAACGTTATACGATGTCCGAGAGTATGCCG	120		

QY	121	GTGTCCTCTTATCAGACCGTTTCCCGCTGGTGAACAGGCGACCGACGCTTTCTGCGAAA	180
Db	121	GTGTCCTCTTATCAGACCGTTTCCCGCTGGTGAACAGGCGACCGACGCTTTCTGCGAAA	180
QY	181	CGCGGAAAAGTGGAGCGCGGATGCGGAGCTGAATACATTCCAAACCGGCTGGCAC	240
Db	181	CGCGGAAAAGTGGAGCGCGGATGCGGAGCTGAATACATTCCAAACCGGCTGGCAC	240
QY	241	AACAACCTGGGGCAACAGTCGTTGCTGATTCGCGCTGCGACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACCTGGGGCAACAGTCGTTGCTGATTCGCGCTGCGACCTCCAGTCTGGCCCTGC	300
QY	301	ACGGCGCTGCGAAAATGTCGGCGGATTAATATCTCGCGCGATCAACTGGGTGCCAGCG	360
Db	301	ACGGCGCTGCGAAAATGTCGGCGGATTAATATCTCGCGCGATCAACTGGGTGCCAGCG	360
QY	361	TGCTGCTGCATGCTGATAGAACGAGCGGCTGGAAGCCTGTAAGCGGCGGTGCAATC	420
Db	361	TGCTGCTGCATGCTGATAGAACGAGCGGCTGGAAGCCTGTAAGCGGCGGTGCAATC	420
QY	421	TTCTCGGCAACCGCTCAGTGGCTGATCAITTAATCTCCGCTGGATGACCAAGGATGCCA	480
Db	421	TTCTCGGCAACCGCTCAGTGGCTGATCAITTAATCTCCGCTGGATGACCAAGGATGCCA	480
QY	481	TTGCTGTGNAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540
Db	481	TTGCTGTGNAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGTTACGGCTGATGACCAAGGATGCCA	600
Db	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGTTACGGCTGATGACCAAGGATGCCA	600
QY	601	TGGTGCATTTGGTCAACAGAAATCGCGCTTTAGCGGGCCCAATAAGTTCTGTCTCGG	660
Db	601	TGGTGCATTTGGTCAACAGAAATCGCGCTTTAGCGGGCCCAATAAGTTCTGTCTCGG	660
QY	661	CGCGTCTGGCTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG	720
Db	661	CGCGTCTGGCTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG	720
QY	721	CGGAACGGGAAGGCGATGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAACGGGAAGGCGATGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCGCATGTTCCCACTCGATGCTGGTGGCAACGATCAGATGGCGCTGGCGCAA	840
Db	781	ATGAGGCGCATGTTCCCACTCGATGCTGGTGGCAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGCGCATTAACCGAGTCCGGGCTGCGGCTGGTGGGATATCTCGGTAGTGGATACG	900
Db	841	TGCGCGCATTAACCGAGTCCGGGCTGCGGCTGGTGGGATATCTCGGTAGTGGATACG	900
QY	901	ACGATACCGAAGACGCTCATGTTATATCCGCGGTTAACCAATCAAAACAGGATTTTC	960
Db	901	ACGATACCGAAGACGCTCATGTTATATCCGCGGTTAACCAATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGCCGCTCTACCTGGTGAAGAAAACCAACCTGGCGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCCGCTCTACCTGGTGAAGAAAACCAACCTGGCGCCCAATA	1080
QY	1081	CGAAACCGGCTCTCCCGCGGCTGGCGGATTCATTAATGAGCTGGACGACAGGTTT	1140
Db	1081	CGAAACCGGCTCTCCCGCGGCTGGCGGATTCATTAATGAGCTGGACGACAGGTTT	1140
QY	1141	CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Db	1141	CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200

QY	1201	GCACAAATTTCTCATGTTTGAACAGCTTATCATCGACTGACCGGTGACCAATCTCTTGGCG	1260
Db	1201	GCACAAATTTCTCATGTTTGAACAGCTTATCATCGACTGACCGGTGACCAATCTCTTGGCG	1260
QY	1261	TCAGGACGCCATCGGAAGCTGTGTATGCTGTGAGGTGCTAAATCATCTGCATAATTCG	1320
Db	1261	TCAGGACGCCATCGGAAGCTGTGTATGCTGTGAGGTGCTAAATCATCTGCATAATTCG	1320
QY	1321	TGTCGCTCAAGCGGCACCTCCCGTTCTGATAATCTTTTTTTCGCGGACATCAACGGTT	1380
Db	1321	TGTCGCTCAAGCGGCACCTCCCGTTCTGATAATCTTTTTTTCGCGGACATCAACGGTT	1380
QY	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAATGTGGGA	1440
Db	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAATGTGGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTAGGTGTTTTACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTAGGTGTTTTACGA	1500
QY	1501	GCATTTCAACAAAGGACCATAGATTTATGAAATCTGAAGAGTAACTGGTAACTCGG	1560
Db	1501	GCATTTCAACAAAGGACCATAGATTTATGAAATCTGAAGAGTAACTGGTAACTCGG	1560
QY	1561	ATTAAACGCGATAAAGCTATACGGTCTCGCTGAAAGTCGGTAAGAAATTCGAGAAGAT	1620
Db	1561	ATTAAACGCGATAAAGCTATACGGTCTCGCTGAAAGTCGGTAAGAAATTCGAGAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTCGAAGAGAAATTCACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTCGAAGAGAAATTCACAGGTT	1680
QY	1681	GGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTTGGTGCTAC	1740
Db	1681	GGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTTGGTGCTAC	1740
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCGGAACAAAGCGTTCCAGGACAAAGTGTAT	1800
Db	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCGGAACAAAGCGTTCCAGGACAAAGTGTAT	1800
QY	1801	CGGTTTACCTGGGATGCCCTGACATTAACACGCAAGCTGATTGCTTACCCGATCGCTGT	1860
Db	1801	CGGTTTACCTGGGATGCCCTGACATTAACACGCAAGCTGATTGCTTACCCGATCGCTGT	1860
QY	1861	GAAAGCTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGAA	1920
Db	1861	GAAAGCTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGAA	1920
QY	1921	GAGATCCCGGCTCGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGGCTCGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACGTAATTCACCTGGCGCTGATTGCTGCTGACGCGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACGTAATTCACCTGGCGCTGATTGCTGCTGACGCGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACCGTGGCGTGGATAACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACCGTGGCGTGGATAACGCTGGCGCGAAAGCG	2100
QY	2101	GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAATGAATGCAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAATGAATGCAGACACCGATTAC	2160
QY	2161	TCCATCGCAGAGTGCCTTTTAAAGGGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAGAGTGCCTTTTAAAGGGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
QY	2221	GCATGGTCCAACTGCAACAGCAACAGCAAGTGAATTAATGCTGTAACGGTACTGCCACCTC	2280
Db	2221	GCATGGTCCAACTGCAACAGCAACAGCAAGTGAATTAATGCTGTAACGGTACTGCCACCTC	2280
QY	2281	AAGGCTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAAGCGCGCAGT	2340

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-15

Perfect score: 3300

Sequence: 1 ccgacacatcgaaatgtgc.....acgacgtccggactacgct 3300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2700	81.8	7475	2	US-08-971-036-1
2	2700	81.8	7475	3	US-09-096-570-1
3	2700	81.8	7475	3	US-09-265-617B-1
C 4	1417.6	43.0	5926	3	US-09-027-169-3
5	1362	41.3	5201	4	US-09-640-882-2
6	1362	41.3	5201	4	US-09-640-882-3
7	1204	36.5	4557	4	US-08-778-717-5
8	1200.8	36.4	3832	1	US-08-148-675A-2
C 9	1199.2	36.3	5248	3	US-08-487-283A-18
C 10	1199.2	36.3	5248	5	PCT-US96-05611A-21
C 11	1199.2	36.3	5312	4	US-10-263-103-35
C 12	1199.2	36.3	5443	2	US-08-929-967-1
C 13	1199.2	36.3	5502	4	US-09-702-705-785
C 14	1199.2	36.3	5502	4	US-09-736-457-785
C 15	1199.2	36.3	5502	4	US-09-614-124B-785
C 16	1199.2	36.3	5502	4	US-09-671-325-785
C 17	1199.2	36.3	5502	4	US-09-589-184-785
C 18	1199.2	36.3	5502	4	US-09-658-824-785
C 19	1199.2	36.3	5616	2	US-08-929-967-3
C 20	1199.2	36.3	5873	4	US-09-695-437A-62
C 21	1199.2	36.3	6353	4	US-09-702-705-784
C 22	1199.2	36.3	6353	4	US-09-736-457-784
C 23	1199.2	36.3	6353	4	US-09-614-124B-784
C 24	1199.2	36.3	6353	4	US-09-671-325-784
C 25	1199.2	36.3	6353	4	US-09-589-184-784
C 26	1199.2	36.3	6353	4	US-09-658-824-784
C 27	1199.2	36.3	6361	3	US-08-646-538-7

C 28	1199.2	36.3	6361	3	US-09-503-222-7
C 29	1199.2	36.3	6727	2	US-08-125-462-2
C 30	1199.2	36.3	6727	2	US-08-891-848-2
C 31	1199.2	36.3	6799	2	US-08-125-462-5
C 32	1199.2	36.3	6799	2	US-08-891-848-5
C 33	1199.2	36.3	7659	3	US-09-128-314-4
C 34	1199.2	36.3	7659	3	US-09-056-556-213
C 35	1199.2	36.3	7676	3	US-08-072-596-208
C 36	1199.2	36.3	7676	4	US-09-072-967-213
C 37	1199.2	36.3	7676	4	US-09-287-849-9
C 38	1199.2	36.3	8031	3	US-09-643-597-254
C 39	1199.2	36.3	8031	4	US-09-480-884A-254
C 40	1199.2	36.3	8031	4	US-09-542-615A-254
C 41	1199.2	36.3	8031	4	US-09-606-421B-254
C 42	1199.2	36.3	8031	4	US-09-476-496A-254
C 43	1199.2	36.3	8031	4	US-09-630-940B-254
C 44	1199.2	36.3	8157	3	US-09-128-314-3
C 45	1199.2	36.3	8501	3	US-08-793-900-1

ALIGNMENTS

RESULT 1
US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8%; Score 2700; DB 2; Length 7475;

Best Local Similarity	100.0%;	Pred. No. 0;
Matches 2700;	Conservative	0; Mismatches
		0; Indels
		0; Gaps

QY	1	CCGACACCATCGAATGGTGTGAAAAACCTTTTCGGGTATGGCATATAGCCGCCGGAAGAGA	60
Db	1	CCGACACCATCGAATGGTGTGAAAAACCTTTTCGGGTATGGCATATAGCCGCCGGAAGAGA	60
QY	61	GTCAATTACAGGTGTGTAATCTGAAACCAAGTAAAGTTATACGATGTGCGACAGTATGCCG	120
Db	61	GTCAATTACAGGTGTGTAATCTGAAACCAAGTAAAGTTATACGATGTGCGACAGTATGCCG	120
QY	121	GTGTCCTTTATCAGACCGTTTCCCGCGTGTGTAAACCAAGCCAGCCAGCCACGTTTCTCGAAAA	180
Db	121	GTGTCCTTTATCAGACCGTTTCCCGCGTGTGTAAACCAAGCCAGCCAGCCACGTTTCTCGAAAA	180
QY	181	CGCGGGAAGAGTGAAGCGCGGATGCGCGAGCTGAAATTAATTCACACCGCTGGCAC	240
Db	181	CGCGGGAAGAGTGAAGCGCGGATGCGCGAGCTGAAATTAATTCACACCGCTGGCAC	240
QY	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCACCTGC	300
Db	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCACCTGC	300
QY	301	ACGCGCGTGCAGAAATTTGTCGCGCGATTAAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCGTGCAGAAATTTGTCGCGCGATTAAATCTCGCGCGATCAACTGGGTGCCAGCG	360
QY	361	TGTTGTTGTCGATGTTAGAACGAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAATC	420
Db	361	TGTTGTTGTCGATGTTAGAACGAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACAGGATGCCA	480
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCCTGCATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540
Db	481	TTGCTGTGGAAGCTGCCTGCATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540
QY	541	CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGGATC	600
Db	541	CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGGATC	600
QY	601	TGTTGCGCATTTGGGTACACAGCAATCGCGTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660
Db	601	TGTTGCGCATTTGGGTACACAGCAATCGCGTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCTGGGATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGCGTCTGGCTGGCTGGGATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTCAACAAAACCATGCAAAATGCTGA	780
Db	721	CGGAAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTCAACAAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATGCTTTCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCAA	840
Db	781	ATGAGGGCATGCTTTCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCAA	840
QY	841	TGCGGCCCATACCGAGTCCGGCTGGGCGTTGGTCGGATATCTCGTATGGGATACG	900
Db	841	TGCGGCCCATACCGAGTCCGGCTGGGCGTTGGTCGGATATCTCGTATGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCATCAACAGGATTTTC	960
QY	961	GCTGCTGGGGCAAAACAGCGTGGACCGTTGTCGAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCTGCTGGGGCAAAACAGCGTGGACCGTTGTCGAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGCGCGTCTCACTGTTGAAAGAAAACCACTCTGGCGCCCAATA	1080

QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db |||||
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db |||||
QY 2221 GCATGGTCAACATCGACACGACGAAAGTGAATTATGGTGAACGTAAGTCTGCGACCTTC 2280
Db |||||
QY 2221 GCATGGTCAACATCGACACGACGAAAGTGAATTATGGTGAACGTAAGTCTGCGACCTTC 2280
Db |||||
QY 2281 AAGGTCACACCATCAACCGTTGCTGGCGTGTGAGCGCAGGATTAACGCCGCCAGT 2340
Db |||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db |||||
QY 2401 GAAAGCGGTTAATAAGACAAACCGTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db |||||
QY 2461 TTGGCGAAAGATCCACGATTTGCCGCCACCATGGGAAACGCCAGAAAGGTGAATCATG 2520
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGATCAACGCC 2580
Db |||||
QY 2581 GCCAGCGTGTGACACTGTCGATGAGCCCTGAAAGAGCGCGACACTAATTCGAGCTCG 2640
Db |||||
QY 2641 AACAAACAAACAATAACAAATCAACAACTCGGATCGAGGGAAGGATTCAGAAATTC 2700
Db |||||

RESULT 2

US-09-096-570-1
; Sequence 1, Application US/09096570
; Patent No. 6018020
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/096,570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522

; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-09-096-570-1

Query Match 81.8%; Score 2700; DB 3; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGTAGCGCCGGAGAGA 60
Db |||||
QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGTAGCGCCGGAGAGA 60
Db |||||
QY 61 GTCATTCAGGGTGGTGAATGTGAAACCAAGTAAAGTATACGATGTCGACAGATATGCCG 120
Db |||||
QY 61 GTCATTCAGGGTGGTGAATGTGAAACCAAGTAAAGTATACGATGTCGACAGATATGCCG 120
Db |||||
QY 121 GTGTCTCTTATCAGACCGTTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
Db |||||
QY 181 CGCGGAAAAAGTGAACGGCGGATGGCGGAGCTGAATTAATTCCTCAACCGCTGGGCAC 240
Db |||||
QY 181 CGCGGAAAAAGTGAACGGCGGATGGCGGAGCTGAATTAATTCCTCAACCGCTGGGCAC 240
Db |||||
QY 241 AACAACTCGCGGAAAAAGTGGTGTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db |||||
QY 241 AACAACTCGCGGAAAAAGTGGTGTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db |||||
QY 301 ACGGCCGTGCAAAATTTCTCGCGCGATTAATTTCTCGCGCGATCAACTGGGTGCCAGCG 360
Db |||||
QY 301 ACGGCCGTGCAAAATTTCTCGCGCGATTAATTTCTCGCGCGATCAACTGGGTGCCAGCG 360
Db |||||
QY 361 TGGTGGTGTGATGTAGAACGAGCGGCTCGAAGCCCTGTAAAGCGGCGGTGCACAATC 420
Db |||||
QY 361 TGGTGGTGTGATGTAGAACGAGCGGCTCGAAGCCCTGTAAAGCGGCGGTGCACAATC 420
Db |||||
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATTAATCTCCGCTGATGACACAGGATGCCA 480
Db |||||
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATTAATCTCCGCTGATGACACAGGATGCCA 480
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTTCTCGCGCGTATTTCTTGATGTCTCTGACCCAGA 540
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTTCTCGCGCGTATTTCTTGATGTCTCTGACCCAGA 540
Db |||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGCACTGGGCGGTGGAGCATC 600
Db |||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGCACTGGGCGGTGGAGCATC 600
Db |||||
QY 601 TGGTGCATTTGGTCCACGACAAATTCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db |||||
QY 601 TGGTGCATTTGGTCCACGACAAATTCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db |||||
QY 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db |||||
QY 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db |||||
QY 721 CCGAACCGGAAAGCGGACCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db |||||
QY 721 CCGAACCGGAAAGCGGACCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db |||||
QY 781 ATGAGGGCATGTTCCCACTCGGATGTTGGTTCACACGATCAGATGCGGTGGCGGCA 840
Db |||||
QY 781 ATGAGGGCATGTTCCCACTCGGATGTTGGTTCACACGATCAGATGCGGTGGCGGCA 840
Db |||||

QY	841	TGCGCGCATTTACGAGTCCGGCTCGCGTTGTCGGGATATCTCGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACGAGTCCGGCTCGCGTTGTCGGGATATCTCGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTTATATCCGCGGTTAACCCACCATCAAAACGAGATTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTTATATCCGCGGTTAACCCACCATCAAAACGAGATTTC	960
QY	961	GCCTGCTGGGGCAAAACAGCGTGACCGCTTGCTGCAACTCTCTCAGGCGCACGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAACAGCGTGACCGCTTGCTGCAACTCTCTCAGGCGCACGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGCTGAAGAAAAACACACCTGGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGCTGAAGAAAAACACACCTGGCGCCCAATA	1080
QY	1081	CGCAAAACCGGCTCTCCCCGGCGGTGGCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Db	1081	CGCAAAACCGGCTCTCCCCGGCGGTGGCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
QY	1141	CCGCACTGGAAGAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGCACTGGAAGAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
QY	1261	TCAGGCAGCCATCGGAAGCTGTGGTAGGTGTGCAGGTGCTAAATCACTGCAATAATTCG	1320
Db	1261	TCAGGCAGCCATCGGAAGCTGTGGTAGGTGTGCAGGTGCTAAATCACTGCAATAATTCG	1320
QY	1321	TGTCGCTCAAGGGCGCATCCCGTCTCGATAATGTTTTTTCGCGCGACATCAACACGGTT	1380
Db	1321	TGTCGCTCAAGGGCGCATCCCGTCTCGATAATGTTTTTTCGCGCGACATCAACACGGTT	1380
QY	1381	CTGGCAAAATCTGAATCAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
Db	1381	CTGGCAAAATCTGAATCAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTACACAGGAAAACAGCCAGTCCGTTTAGTGTGTTTTACGA	1500
Db	1441	ATTGTGAGCGGATAACAATTTACACAGGAAAACAGCCAGTCCGTTTAGTGTGTTTTACGA	1500
QY	1501	GCATTTACCAACAGAGCAATAGATTATGAAAACCTGAAGAGTAACTGGTAAATCTCG	1560
Db	1501	GCATTTACCAACAGAGCAATAGATTATGAAAACCTGAAGAGTAACTGGTAAATCTCG	1560
QY	1561	ATTAAACGGCGATAAAGGCTAACGGTCTCGCTGAAGTCCGTTAGAAATTCGAGAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTAACGGTCTCGCTGAAGTCCGTTAGAAATTCGAGAAGAT	1620
QY	1621	ACCGGAATTAAGTTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT	1680
Db	1621	ACCGGAATTAAGTTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT	1680
QY	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTGGTGGCTAC	1740
Db	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACACAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATGCGGTACGTTACACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCGGTACGTTACACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
QY	1861	GAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCGGCAAAAACTGGGAA	1920
Db	1861	GAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCGGCAAAAACTGGGAA	1920

Qy	1921	GAGATCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGATGTTCAAC	1980
	1981	CTGCAAGAACCGTACTTCACCTGGCGCTGATTCGCTCAGCGGGGGTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACCTGGCGCTGATTCGCTCAGCGGGGGTATGCGTTCAAG	2040
Qy	2041	TATGAAACCGCAAGTACGACATTAAAGACGCTGGCGCTGGATACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAACCGCAAGTACGACATTAAAGACGCTGGCGCTGGATACGCTGGCGCGAAAGCG	2100
Qy	2101	GGTCTGACCTTCCTGGTTGACCTGATTAAAAAACAAACACATGAATGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCCTGGTTGACCTGATTAAAAAACAAACACATGAATGAGACACCGATTAC	2160
Qy	2161	TCCATCCGAGAAGCTGCTTTAATAAAGGCGAACACGCGATGCCATCAACGCGCCCGTGG	2220
Db	2161	TCCATCCGAGAAGCTGCTTTAATAAAGGCGAACACGCGATGCCATCAACGCGCCCGTGG	2220
Qy	2221	GCATGCTCCAACTCGACACACGACAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGCTCCAACTCGACACACGACAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGTCACATCCATAACCGCTTCGTTGGCGTGTGAGCGCAGGATTAACGCGCCCGCAGT	2340
Db	2281	AAGGTCACATCCATAACCGCTTCGTTGGCGTGTGAGCGCAGGATTAACGCGCCCGCAGT	2340
Qy	2341	CCGAAACAAAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAAACAAAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGTCTG	2400
Qy	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAAGATCCACGTAATTGCGGCACCATGGAAAAACGCCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTAATTGCGGCACCATGGAAAAACGCCCAGAAAGGTGAATCATG	2520
Qy	2521	CCGAAACATCCGCGAGATGCCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAAACATCCGCGAGATGCCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTCGTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTCGTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAAACAAACAAATAACAATAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC	2700
Db	2641	AACAAACAAACAAATAACAATAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC	2700

RESULT 3

```

US-09-265-617B-1
; Sequence 1, Application US/09265617B
; Patent No. 6372883
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R.
; APPLICANT: Hurst, David N.
; APPLICANT: Jones, Philip S.
; APPLICANT: Kay, Paul B.
; APPLICANT: Raynham, Tony M.
; APPLICANT: Wilson, Francis X.
; TITLE OF INVENTION: Antiviral Medicaments
; FILE REFERENCE: 20052 antiviral medicaments
; CURRENT APPLICATION NUMBER: US/09/265,617B
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: GB9606815.8
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

```

; LENGTH: 7475									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: Vector and									
; OTHER INFORMATION: Gene Fragments									
US-09-265-617B-1									
Query Match 81.8%; Score 2700; DB 3; Length 7475;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CCGACACCATGAAATGGTGC	AAACCTTTCCGGTATGGCATG	ATAGCCCGGAAGAGA	60				
DB	1	CCGACACCATGAAATGGTGC	AAACCTTTCCGGTATGGCATG	ATAGCCCGGAAGAGA	60				
QY	61	GTCAATTCAGGGTGGTGAATG	TGAACACAGTAACTATACGATG	TCGCAGAGTATGCGG	120				
DB	61	GTCAATTCAGGGTGGTGAATG	TGAACACAGTAACTATACGATG	TCGCAGAGTATGCGG	120				
QY	121	GTGTCCTTATCAGACCGT	TTCCCGGTGGTGAACACAGGCCAGCCAGCGTTTCTGCGAAAA	180					
DB	121	GTGTCCTTATCAGACCGT	TTCCCGGTGGTGAACACAGGCCAGCCAGCGTTTCTGCGAAAA	180					
QY	181	CGCGGAAAAAGTGAACGGCGG	ATGCGGAGCTGAATTACATTC	CCCAACCGCGTGGCAC	240				
DB	181	CGCGGAAAAAGTGAACGGCGG	ATGCGGAGCTGAATTACATTC	CCCAACCGCGTGGCAC	240				
QY	241	AACAACTCGCGGGCAAA	CAGTCTGTTGCTGATGGCGTTGGCCACTCCAGTCTGGGCCCTGC	300					
DB	241	AACAACTCGCGGGCAAA	CAGTCTGTTGCTGATGGCGTTGGCCACTCCAGTCTGGGCCCTGC	300					
QY	301	ACGCGCGTTCGCAAAAT	GTTCGGCGGATTAATCTTCGGCCGATCAACTGGGTGCCAGCG	360					
DB	301	ACGCGCGTTCGCAAAAT	GTTCGGCGGATTAATCTTCGGCCGATCAACTGGGTGCCAGCG	360					
QY	361	TGGTGGTTCGATGTAGAAC	AGAGGGCGTGCAGAGCTGTAAAGCGGCGTGCACAATC	420					
DB	361	TGGTGGTTCGATGTAGAAC	AGAGGGCGTGCAGAGCTGTAAAGCGGCGTGCACAATC	420					
QY	421	TTCTCGCGCAACGCGT	CAGTGGGCTGATCATTAATCTATCCGCTGGATGACACAGGATGCCA	480					
DB	421	TTCTCGCGCAACGCGT	CAGTGGGCTGATCATTAATCTATCCGCTGGATGACACAGGATGCCA	480					
QY	481	TTGCTGTGGAAGCTGC	TGCACATAATGTTTCGGCGGTTATTTCTTGATGTTCTCTGACCCAGA	540					
DB	481	TTGCTGTGGAAGCTGC	TGCACATAATGTTTCGGCGGTTATTTCTTGATGTTCTCTGACCCAGA	540					
QY	541	CACCCATCAACAGTAT	TATTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATC	600					
DB	541	CACCCATCAACAGTAT	TATTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATC	600					
QY	601	TGGTCGCATTGGGT	CACAGCAAAATCGGCTGTAGCGGGCCCATTAAGTTCGTCTCGG	660					
DB	601	TGGTCGCATTGGGT	CACAGCAAAATCGGCTGTAGCGGGCCCATTAAGTTCGTCTCGG	660					
QY	661	CGCGTCTGCGTCTG	CTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG	720					
DB	661	CGCGTCTGCGTCTG	CTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG	720					
QY	721	CGGAACGGGAAGCG	CACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA	780					
DB	721	CGGAACGGGAAGCG	CACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA	780					
QY	781	ATGAGGGCATCGT	CCCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGGCGCAA	840					
DB	781	ATGAGGGCATCGT	CCCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGGCGCAA	840					
QY	841	TGCGGCCATTACCG	AGTCCGGGCTCGCGGTTGGTGGGATATCTCGGTAGTGGGATAGC	900					
DB	841	TGCGGCCATTACCG	AGTCCGGGCTCGCGGTTGGTGGGATATCTCGGTAGTGGGATAGC	900					
QY	901	ACGATACCGAAGAC	AGCTCATGTTATATCCCGCGGTTAAACCAACCATCAAAAGGATTTTC	960					

DB	901	ACGATACCGAAGACAGCT	CACTGTTATATCCCGCGTAAACCAACCATCAAAAGGATTTTC	960					
QY	961	GCCTGCTGGGCAAAAC	CCAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020					
DB	961	GCCTGCTGGGCAAAAC	CCAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020					
QY	1021	AGGGCAATCAGCTGT	TGCCCGTCTCACTGGTGAAGAAAGAAACCAACCTTGGCGGCCAATA	1080					
DB	1021	AGGGCAATCAGCTGT	TGCCCGTCTCACTGGTGAAGAAAGAAACCAACCTTGGCGGCCAATA	1080					
QY	1081	CGAAAAACCGCTCT	CCCCCGCGCTTGGCGATTCATTAATGCAGCTGCACGACAGGTTT	1140					
DB	1081	CGAAAAACCGCTCT	CCCCCGCGCTTGGCGATTCATTAATGCAGCTGCACGACAGGTTT	1140					
QY	1141	CCCGACTGGAAGCGG	GCAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG	1200					
DB	1141	CCCGACTGGAAGCGG	GCAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG	1200					
QY	1201	GCACAAATCTCAT	TTTGACAGCTTATCATCGCTGACCGGTGCACCAATGCTTCTGGCG	1260					
DB	1201	GCACAAATCTCAT	TTTGACAGCTTATCATCGCTGACCGGTGCACCAATGCTTCTGGCG	1260					
QY	1261	TCAGCAGCCATCG	GAAGCTGTGTATGCTGTGAGCTCGTAAATCACTGCATAAATCG	1320					
DB	1261	TCAGCAGCCATCG	GAAGCTGTGTATGCTGTGAGCTCGTAAATCACTGCATAAATCG	1320					
QY	1321	TGTGCTCAAGCG	GCACCTCCCGTCTTGATTAATGTTTTGCGCGGACATCAATACGGTT	1380					
DB	1321	TGTGCTCAAGCG	GCACCTCCCGTCTTGATTAATGTTTTGCGCGGACATCAATACGGTT	1380					
QY	1381	CTGGCAAAATAT	CTGAAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTGTGGA	1440					
DB	1381	CTGGCAAAATAT	CTGAAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTGTGGA	1440					
QY	1441	ATTGTGACGGAT	AACTTTTACACAGAAACAGCAGTCCGTTTAGTGTTCACGA	1500					
DB	1441	ATTGTGACGGAT	AACTTTTACACAGAAACAGCAGTCCGTTTAGTGTTCACGA	1500					
QY	1501	GCACCTTCAAC	CAAGGACCATAGATTATGAAGAGGTAATCTGGTAATCTGG	1560					
DB	1501	GCACCTTCAAC	CAAGGACCATAGATTATGAAGAGGTAATCTGGTAATCTGG	1560					
QY	1561	ATTAACCGCGAT	TAAGGCTATAACCGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620					
DB	1561	ATTAACCGCGAT	TAAGGCTATAACCGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620					
QY	1621	ACCGGAATTAAG	TACACCGTTGAGCATCCGATAACTGGAGAGAAATTCACAGGTT	1680					
DB	1621	ACCGGAATTAAG	TACACCGTTGAGCATCCGATAACTGGAGAGAAATTCACAGGTT	1680					
QY	1681	GGCGCAACTGG	CGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTGGTGGCTAC	1740					
DB	1681	GGCGCAACTGG	CGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTGGTGGCTAC	1740					
QY	1741	GCTCAATCTG	CGCTTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT	1800					
DB	1741	GCTCAATCTG	CGCTTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT	1800					
QY	1801	CGGTTTACCT	GGGATGCGGTAGTTTACAAACCGCAAGCTGTTTACCGATCCGCTGT	1860					
DB	1801	CGGTTTACCT	GGGATGCGGTAGTTTACAAACCGCAAGCTGTTTACCGATCCGCTGT	1860					
QY	1861	GAAGCGTTAT	CGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920					
DB	1861	GAAGCGTTAT	CGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920					
QY	1921	GAGATCCCG	CGCTGGATTAAGAACTGAAGCGAAGGTAAGAGCGCTGATGTTCAAC	1980					
DB	1921	GAGATCCCG	CGCTGGATTAAGAACTGAAGCGAAGGTAAGAGCGCTGATGTTCAAC	1980					
QY	1981	CTGCAAGAAC	CGTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATCGGTTCAAG	2040					

Db 1981 CTGCAAGAACCTGACTTCACTGCGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACCTGGCGGTGGATACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACCTGGCGGTGGATACGCTGGCGGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACAACATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACAACATGAATGACACACCGATTAC 2160
Qy 2161 TCCATCCGAGAGCTGCTTAAATAGAGCGCAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTAAATAGAGCGCAACAGCGATGACCATCAACGGCCGCTGG 2220
Qy 2221 GCATGGTCCCAATCGACACCAAGTGAATTAATGTTGTAACCGTACTCCCGACCTTC 2280
Db 2221 GCATGGTCCCAATCGACACCAAGTGAATTAATGTTGTAACCGTACTCCCGACCTTC 2280
Qy 2281 AAGGTCACACCATCAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCATCAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACCAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACCAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAACGGTAAATAAGACCAACCGCTGGTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2460
Db 2401 GAACGGTAAATAAGACCAACCGCTGGTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2460
Qy 2461 TTGCGGAAAGTCAAGTATTTGCGCGCACCATGAAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGCGGAAAGTCAAGTATTTGCGCGCACCATGAAACCGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGCTCGCTTTCTGTTGATGCGTGTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCGCGAGATGCTCGCTTTCTGTTGATGCGTGTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACCAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTCAGAAATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTCAGAAATTC 2700

RESULT 4

US-09-027-169-3/c
; Sequence 3, Application US/09027169
; Patent No. 6420524
; GENERAL INFORMATION:
; APPLICANT: CRAIG, NANCY L
; TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
; TITLE OF INVENTION: ATP-DEPENDENT TRANSCRIPTION PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Anne Brown (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave.
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,169
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Brown, Anne
; REGISTRATION NUMBER: 36,463
; REFERENCE/DOCKET NUMBER: 5789-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2205
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "pEM delta R.adj to 1"
; US-09-027-169-3

Query Match 43.0%; Score 1417.6; DB 3; Length 5926;

Best Local Similarity 98.6%; Pred. No. 0;

Matches 1455; Conservative 0; Mismatches 9; Indels 12; Gaps 2;

Qy	3	GACACCATCGAATGGTGCRAAACCTTTTCGGGTATGCGCATGATAGCGCCGGAAGAGAGT	62
Db	1564	GACACCATCGAATGGTGCRAAACCTTTTCGGGTATGCGCATGATAGCGCCGGAAGAGAGT	1505
Qy	63	CAATTTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTGCGAGAGTATGCCGGT	122
Db	1504	CAATTTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTGCGAGAGTATGCCGGT	1445
Qy	123	GTCTCTTATCAGACCGTTTCCCGGTGGTGAAACAGGCCAGCCAGCTTTCTGGAAAAACG	182
Db	1444	GTCTCTTATCAGACCGTTTCCCGGTGGTGAAACAGGCCAGCCAGCTTTCTGGAAAAACG	1385
Qy	183	CGGGAAAAAGTGAAGCGGCGATGGCGGAGCTCAATTACATTCCTCCACCGGTGGCAAA	242
Db	1384	CGGGAAAAAGTGAAGCGGCGATGGCGGAGCTCAATTACATTCCTCCACCGGTGGCAAA	1325
Qy	243	CAACTGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGGCCCTGCAC	302
Db	1324	CAACTGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGGCCCTGCAC	1265
Qy	303	GCSCGTGCGAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCGAGCGT	362
Db	1264	GCSCGTGCGAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCGAGCGT	1205
Qy	363	GTGCTGTGATGTAGAAAGCAAGCGCGCTGAAAGCGCTGTAAGCGCGGTGCAATCTT	422
Db	1204	GTGCTGTGATGTAGAAAGCAAGCGCGCTGAAAGCGCTGTAAGCGCGGTGCAATCTT	1145
Qy	423	CTCGCGCAACGCTCAGTGGGCTGATCATTAACATCTCCGCTGGATGACCAAGGATGCCATT	482
Db	1144	CTCGCGCAACGCTCAGTGGGCTGATCATTAACATCTCCGCTGGATGACCAAGGATGCCATT	1085
Qy	483	GCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTTCTGTGACGACACA	542
Db	1084	GCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTTCTGTGACGACACA	1025
Qy	543	CCCATCAACAGTATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATCTG	602
Db	1024	CCCATCAACAGTATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATCTG	965
Qy	603	GTGCGATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTGGCG	662
Db	964	GTGCGATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTGGCG	905
Qy	663	CGTCTGGTCTGGCTGGCGTAATATCTCCTACCTCGCAATCAAAATTCAGCGGATAGCG	722
Db	904	CGTCTGGTCTGGCTGGCGTAATATCTCCTACCTCGCAATCAAAATTCAGCGGATAGCG	845
Qy	723	GAAACGGGAAGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGAAT	782
Db	844	GAAACGGGAAGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGAAT	785
Qy	783	GAGGGCATCGTTCCCACTCGCATGCTGGTGTGCGCAACGATCGATGGCGTGGCGCAATG	842

Db 784 GAGGGCATCGTTCCCACTGCGATGCTGGTTGGTCCAAACGATCAGATGGCGTGGGCCAATG 725
Qy 843 CGCGCATTTACCGAGTCCGGGCTGGCGCTGGTGGGATATCTCGGTAGTGGATACGAC 902
Db 724 CGCGCATTTACCGAGTCCGGGCTGGCGCTGGTGGGATATCTCGGTAGTGGATACGAC 665
Qy 903 GATACCGAAGCAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGAGTTTTCG 962
Db 664 GATACCGAAGCAGCTCATGTTATATCCCGCGTTCAACCAACCATCAACAGAGTTTTCG 605
Qy 963 CTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 604 CTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 545
Qy 1023 GCGAATCAGCTGTTCCCGCTCTCACTGGTGGAAAGAAACCAACCGCTGGCGCCCAATACG 1082
Db 544 GCGAATCAGCTGTTCCCGCTCTCACTGGTGGAAAGAAACCAACCGCTGGCGCCCAATACG 485
Qy 1083 CAAACCGGCTCTCCCGCGGTGGCGGATTCATTAATGACGCTGGCAGCAGAGTTTCC 1142
Db 484 CAAACCGGCTCTCCCGCGGTGGCGGATTCATTAATGACGCTGGCAGCAGAGTTTCC 425
Qy 1143 CGACTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGAGTTAGCTCACTCATTAGGC 1202
Db 424 CGACTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGAGTTAGC-----GC 376
Qy 1203 ACAATTTCTCATGTTTGACAGCTTTATCATCGACTGCACGGTGCAACCAATGTTCTGGCGTC 1262
Db 375 GAAATTGATCTGTTTGACAGCTTTATCATCGACTGCACGGTGCAACCAATGTTCTGGCGTC 316
Qy 1263 AGGAGCCATCGGAAGCTGTGATGGTGTGCAGGTCGTAATCACTGCATAATTCGTG 1322
Db 315 AGGAGCCATCGGAAGCTGTGATGGTGTGCAGGTCGTAATCACTGCATAATTCGTG 256
Qy 1323 TCGCTCAAGGCGCACTCCGCTTCGATAATGTTTTTGGCGCGACATCAACGGTTCT 1382
Db 255 TCGCTCAAGGCGCACTCCGCTTCGATAATGTTTTTGGCGCGACATCAACGGTTCT 196
Qy 1383 GCGAAATATTCGAAATGAGCTGTTGACAAATTAATCAT--CGGCTCGTATAATGTTGGAA 1441
Db 195 GCGAAATATTCGAAATGAGCTGTTGACAAATTAATCATCAATCCGGCTCGTATAATGTTGGAA 136
Qy 1442 TTGTGAGCGGATPAACAATTTACAGAGGAACAGCC 1477
Db 135 TTGTGAGCGGATPAACAATTTACAGAGGAACAGAC 100

RESULT 5

US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE2
US-09-640-882-2

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
Qy 3 GACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCCGAGAGAGT 62
Db 866 GACACCATCGAATGGCGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCCGAGAGAGT 925
Qy 63 CAATTACAGGTGGTGAATGTGAAACCCAGTAACTGATACGATGTGCGAGAGATATGCCGGT 122
Db 926 CAATTACAGGTGGTGAATGTGAAACCCAGTAACTGATACGATGTGCGAGAGATATGCCGGT 985
Qy 123 GTCTCTTATCAGACGGTTTCCCGGTGGTGAACAGCGCCAGCAGCTTTCTGCGAAGACG 182
Db 986 GTCTCTTATCAGACGGTTTCCCGGTGGTGAACAGCGCCAGCAGCTTTCTGCGAAGACG 1045
Qy 183 CGGGAAGGTGGAAGCGGCGATGGCGAGCTGAATTAATTCCTCCAAACCGGTGGCAAA 242
Db 1046 CGGGAAGGTGGAAGCGGCGATGGCGAGCTGAATTAATTCCTCCAAACCGGTGGCAAA 1105
Qy 243 CAATGCGGCGCAACAGTGGTCTGATTCGCGGTGGCCACTCCAGTCTGGCCCTGCAC 302
Db 1106 CAATGCGGCGCAACAGTGGTCTGATTCGCGGTGGCCACTCCAGTCTGGCCCTGCAC 1165
Qy 303 CGCGCTGCGAAATTTGTCGCGCGATTAATTCCTCGCGCGATCAACTGGGTGCCAGCGTG 362
Db 1166 CGCGCTGCGAAATTTGTCGCGCGATTAATTCCTCGCGCGATCAACTGGGTGCCAGCGTG 1225
Qy 363 GTGGTGTGATGATGAAGCGGCGTGAAGCGCTGTAAGCGGCGGTGGCAATCTT 422
Db 1226 GTGGTGTGATGATGAAGCGGCGTGAAGCGCTGTAAGCGGCGGTGGCAATCTT 1285
Qy 423 CTCGCGCAACCGTCACTGGGCTGATCAATTAATTCCTCGCTGGATGACAGGATGCCATT 482
Db 1286 CTCGCGCAACCGTCACTGGGCTGATCAATTAATTCCTCGCTGGATGACAGGATGCCATT 1345
Qy 483 GCTGTGGAAGCTGCTGCACATAATGTTCCGGGCTTAATTTCTTGATGCTCTGACAGACA 542
Db 1346 GCTGTGGAAGCTGCTGCACATAATGTTCCGGGCTTAATTTCTTGATGCTCTGACAGACA 1405
Qy 543 CCCATCAACAGTATTTATTTCTCCATGAAGCGTACGCGACTGGGCGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTTATTTCTCCATGAAGCGTACGCGACTGGGCGTGGAGCATCTG 1465
Qy 603 GTCGCAATGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGGG 662
Db 1466 GTCGCAATGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGGG 1525
Qy 663 GGTCTGCGTCTGGCTGGCGTGAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 1526 GGTCTGCGTCTGGCTGGCGTGAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAAAGGGAAGCGACTGGAGTGCATGTCGCGTTTCAACAAACCATCAAAATGCTGAAT 782
Db 1586 GAAAGGGAAGCGACTGGAGTGCATGTCGCGTTTCAACAAACCATCAAAATGCTGAAT 1645
Qy 783 GAGGCGATCGTTCCCACTGCGATGCTGTTGCGCAACGATCAGATGGCGTGGCGCAATG 842
Db 1646 GAGGCGATCGTTCCCACTGCGATGCTGTTGCGCAACGATCAGATGGCGTGGCGCAATG 1705
Qy 843 CGCGCATTTACGAGTCCGGCTGGCGGATATCTCGGTAGTGGGATACGAC 902
Db 1706 CGCGCATTTACGAGTCCGGCTGGCGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGCAGCTCATGTTATATCCCGCGTTTAACCAACCATCAACAGAGTTTTCG 962
Db 1766 GATACCGAAGCAGCTCATGTTATATCCCGCGTTTAACCAACCATCAACAGAGTTTTCG 1825
Qy 963 CTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 1826 CTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 1885
Qy 1023 GCGAATCAGCTGTTGCGCGTCTCACTGTTGAAAGAAACCAACCGCTGGCGCCCAATACG 1082

```
Db 1886 GGCATTCAGCTGTGCGCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATAGC 1945
Qy 1083 CAAACCGCCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTTCC 1142
Db 1946 CAAACCGCCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTTCC 2005
Qy 1143 CGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTATTAGGC 1202
Db 2006 CGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAG--CGCGAATTGATC 2063
Qy 1203 ACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGACCGTGACCAATGCTTCTGGCGTC 1262
Db 2064 TGAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGACCGTGACCAATGCTTCTGGCGTC 2123
Qy 1263 AGGAGCCATCGGAAGCTGTGTGTATGGCTGTGCAGTCTGTAATCACTGCATAATTCTGTG 1322
Db 2124 AGGAGCCATCGGAAGCTGTGTGTATGGCTGTGCAGTCTGTAATCACTGCATAATTCTGTG 2183
Qy 1323 TGGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTTCGCGCGACATCATACGGTTCT 1382
Db 2184 TGGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTTCGCGCGACATCATACGGTTCT 2243
Qy 1383 GSCAAATATTCT 1394
Db 2244 GSCAAATATTCT 2255

RESULT 6
US-09-640-882-3
; Sequence 3, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
US-09-640-882-3

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 3 GACACCATCGAATCGTGGCAACCTTTTCGGGTATGGCATGATAGCCCGGAGAGAGT 62
Db 866 GACACCATCGAATCGTGGCAACCTTTTCGGGTATGGCATGATAGCCCGGAGAGAGT 925
Qy 63 CAATTCAGGTTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCGT 122
Db 926 CAATTCAGGTTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCGT 985
Qy 123 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCAGCTTTCTGGGAAACG 182
Db 986 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCAGCTTTCTGGGAAACG 1045
Qy 183 CGGGAAGAGTGGAGCGGCGATGGCGAGCTGAAATTAATCCCAACCCGCTGGCACA 242
Db 1046 CGGGAAGAGTGGAGCGGCGATGGCGAGCTGAAATTAATCCCAACCCGCTGGCACA 1105
```

```
Qy 243 CAACCTGGCGGCAACAGTCGTCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 302
Db 1106 CAACCTGGCGGCAACAGTCGTCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 1165
Qy 303 GCGCCGTCGCAAAATTTGTTCGGCGGATTTAAATCTCGCGCCGATCAATCGGGTGCCAGGCTG 362
Db 1166 GCGCCGTCGCAAAATTTGTTCGGCGGATTTAAATCTCGCGCCGATCAATCGGGTGCCAGGCTG 1225
Qy 363 GTGTGTTCGATGTAGAACGAGCGGCTGGAAGCCCTGTAAAGCGCGGTGCAACAATCTT 422
Db 1226 GTGTGTTCGATGTAGAACGAGCGGCTGGAAGCCCTGTAAAGCGCGGTGCAACAATCTT 1285
Qy 423 CTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGGATGTCATT 482
Db 1286 CTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGGATGTCATT 1345
Qy 483 GTGTGGAAGTGCCTGCATTAATGTTCCGGCGTTATTTCTTGTATGTTCTTGACAGACA 542
Db 1346 GTGTGGAAGTGCCTGCATTAATGTTCCGGCGTTATTTCTTGTATGTTCTTGACAGACA 1405
Qy 543 CCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATCTG 1465
Qy 603 GTCGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAACTTCTGTCGCGG 662
Db 1466 GTCGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAACTTCTGTCGCGG 1525
Qy 663 GTCGCTGCTGGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 1526 GTCGCTGCTGGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAACGGGAAGCGCACTGGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGAAT 782
Db 1586 GAACGGGAAGCGCACTGGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGAAT 1645
Qy 783 GAGGGCATCGTCCCACTGCGATGCTGTTTGCACAGCATCAGATGCGCTGGCGCGCAATG 842
Db 1646 GAGGGCATCGTCCCACTGCGATGCTGTTTGCACAGCATCAGATGCGCTGGCGCGCAATG 1705
Qy 843 CGCGCATTAACGAGTCCGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATACGAC 902
Db 1706 CGCGCATTAACGAGTCCGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCAACCATCAAAACAGGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCAACCATCAAAACAGGATTTTCG 1825
Qy 963 CTGCTGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 1826 CTGCTGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGAAG 1885
Qy 1023 GGCATACAGTGTGCGCGCTCACTGTTGAAAGAAACCAACCCCTGGCGCCCAATAGC 1082
Db 1886 GGCATACAGTGTGCGCGCTCACTGTTGAAAGAAACCAACCCCTGGCGCCCAATAGC 1945
Qy 1083 CAAACCGCCTCTCCCGCGCGTTGGCGGATTCATTAAATGAGCTGGCAGCAGAGTTTTC 1142
Db 1946 CAAACCGCCTCTCCCGCGCGTTGGCGGATTCATTAAATGAGCTGGCAGCAGAGTTTTC 2005
Qy 1143 CGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTATTAGGC 1202
Db 2006 CGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAG--CGCGAATTGATC 2063
Qy 1203 ACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGACCGTGACCAATGCTTCTGGCGTC 1262
Db 2064 TGAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGACCGTGACCAATGCTTCTGGCGTC 2123
Qy 1263 AGGAGCCATCGGAAGCTGTGTATGGCTGTGCAGTCTGTAATCACTGCATAATTCTGTG 1322
Db 2124 AGGAGCCATCGGAAGCTGTGTATGGCTGTGCAGTCTGTAATCACTGCATAATTCTGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATACGGTTCT 1382
```

|||||
Db 2184 TCGCTCAAGCGCACTCCGTTCTGGATAATGTTTTTGGCCGACATCAACGGTTCT 2243
QY 1383 GCGAAATATTCT 1394
|||||
Db 2244 GCGAAATATTCT 2255
|||||
RESULT 7
US-08-778-717-5
; Sequence 5, Application US/08778717
; Patent No. 6602689
; GENERAL INFORMATION:
; APPLICANT: USNO, EIICHI
; APPLICANT: NOBUYUKI, FUJII
; APPLICANT: OKADA, NASHISA
; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN
; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR
; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEURSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/778,717
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 352225/1995
; FILING DATE: 28-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2084-031-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4557 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; ORIGINAL SOURCE:
; ORGANISM: E. COLI
; STRAIN: BL21 (DE3)
; PUBLICATION INFORMATION:
; AUTHORS: NOBUYUKI FUJII ET AL,
; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED
; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF
; TITLE: EXPRESSING SAID FUSED PROTEIN
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557
US-08-778-717-5
Query Match 36.5%; Score 1204; DB 4; Length 4557;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACCATCAATGGTCAAAACCTTTCGGGGTATGGCATATAGCCCGGAGAGA 60
|||||
Db 2563 CCGACACCATCAATGGTCAAAACCTTTCGGGGTATGGCATATAGCCCGGAGAGA 2622

QY 61 GTCAATTTCAGGGTGGTGAATGTAAGCAAGTAAAGTTATACGATGTCGACAGTATGCGG 120
Db 2623 GTCAATTTCAGGGTGGTGAATGTAAGCAAGTAAAGTTATACGATGTCGACAGTATGCGG 2682
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGCCAGCTTTCTGCGAAAA 180
Db 2683 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGCCAGCTTTCTGCGAAAA 2742
QY 181 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGCTGGGCAC 240
Db 2743 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGCTGGGCAC 2802
QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATGTTGGCTTGCACCTCCAGTCTGCGCCCTGC 300
Db 2803 AACAACTGGCGGCAAAACAGTCGTTGCTGATGTTGGCTTGCACCTCCAGTCTGCGCCCTGC 2862
QY 301 ACGCGCCCTCGCAAAATTCGCGCGGATTAATCTCGCGCGGATCAATCGGGGTGCCAGCG 360
Db 2863 ACGCGCGTTCGCAAAATTCGCGCGGATTAATCTCGCGCGGATCAATCGGGGTGCCAGCG 2922
QY 361 TGGTGGTTCGATGGTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCAAAATC 420
Db 2923 TGGTGGTTCGATGGTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCAAAATC 2982
QY 421 TTCTCGCGCAACGGTCAAGTGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
Db 2983 TTCTCGCGCAACGGTCAAGTGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3042
QY 481 TTGTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA 540
Db 3043 TTGTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA 3102
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGGGTGGAGCATC 600
Db 3103 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGGGTGGAGCATC 3162
QY 601 TGGTGGCATTTGGGTCAACAGCAATTCGCGCTGTTAGCGGGGCCCATTAAGTCTGCTCGG 660
Db 3163 TGGTGGCATTTGGGTCAACAGCAATTCGCGCTGTTAGCGGGGCCCATTAAGTCTGCTCGG 3222
QY 661 CGCGTCTCGCTCTGGCTGGGATAAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 3223 CGCGTCTCGCTCTGGCTGGGATAAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 3282
QY 721 CGGAAACGGGAAGCGCACTGGAGTGCCATGTCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 3283 CGGAAACGGGAAGCGCACTGGAGTGCCATGTCGGTTTTTCAACAAACCAATGCAAAATGCTGA 3342
QY 781 ATGAGGGCATCGTTCCCACTGGGATGCTGTTGCCAAGATCAGATGGCGCTGGGGGCA 840
Db 3343 ATGAGGGCATCGTTCCCACTGGGATGCTGTTGCCAAGATCAGATGGCGCTGGGGGCA 3402
QY 841 TCGCGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db 3403 TCGCGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 3462
QY 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC 960
Db 3463 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC 3522
QY 961 GCCTCTGGGGCAAAACAGCGTGAGCCGCTTGTGCTCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 3523 GCCTCTGGGGCAAAACAGCGTGAGCCGCTTGTGCTCAACTCTCTCAGGCGCCAGGCGGTGA 3582
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATA 1080
Db 3583 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATA 3642
QY 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAAATGCAGCTGGCACACAGGTTT 1140
Db 3643 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAAATGCAGCTGGCACACAGGTTT 3702
QY 1141 CCCGACTGGAAGCGGGGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTACTCATTAG 1200

|||||
Db 3703 CCCGACTGAAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTAGTCACTCATTAG 3762
Qy 1201 GCAC 1204
Db 3763 GCAC 3766
|||||
RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerita, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3832 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-148-675A-2
Query Match 36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGATGGTGCAGAAACCTTCCGCGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 3 CCGACACCATCGAATGGCGCAAAACCTTTCGCGTATGGCATGATAGCGCCCGGAAGAGA 62
Qy 61 GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTTATACGATGTCGCGAGAGTATGCCG 120
Db 63 GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTTATACGATGTCGCGAGAGTATGCCG 122
Qy 121 GTGTCTCTTATCAGACCGTTTCCGCGTGTGTAACAGCGCCAGCCAGTTTCTCGCAAAA 180
Db 123 GTGTCTCTTATCAGACCGTTTCCGCGTGTGTAACAGCGCCAGTTTCTCGCAAAA 182
Qy 181 CGCGGGAAGAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCCTCAACCGGTGGCAC 240
Db 183 CGCGGGAAGAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCCTCAACCGGTGGCAC 242
Qy 241 AACAACTGGCGGGCAAAACAGTCGTTGTGATTTGGCGTTGGCACCTCCAGTCTCGGCCCTGC 300
|||||

Db 243 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCGCCCTGC 302
Qy 301 ACGCCCGTTCGCAAAATTTGTCCGCGCATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 303 ACGCCCGTTCGCAAAATTTGTCCGCGCATTAATCTCGCGCGATCAACTGGGTGCCAGCG 362
Qy 361 TGGTGGTGTGATGGTAGAACGAAAGCGCGTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db 363 TGGTGGTGTGATGGTAGAACGAAAGCGCGTCGAAGCCTGTAAAGCGCGGTGCACAATC 422
Qy 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACACAGGATGCCA 480
Db 423 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACACAGGATGCCA 482
Qy 481 TTGCTGTGGAGCTGCGCTGCACTAAATGTTCCGGGGTTATTTCTTGATGTCTTGACGAGA 540
Db 483 TTGCTGTGGAGCTGCGCTGCACTAAATGTTCCGGGGTTATTTCTTGATGTCTTGACGAGA 542
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 543 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 602
Qy 601 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 603 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 662
Qy 661 CGGCTCTCGCTGCGCTGGCTGGCGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 663 CGGCTCTCGCTGCGCTGGCTGGCGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 722
Qy 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCGGTTCCTCAACAAACCATCAAAATGCTGA 780
Db 723 CGGAACGGGAAGCGCACTGGAGTGCCATGTCGGTTCCTCAACAAACCATCAAAATGCTGA 782
Qy 781 ATGAGGGCATCGTTTCCCATCGCATGTCGATGTCGCAACGATCAGATGGCGCTGGCGCAA 840
Db 783 ATGAGGGCATCGTTTCCCATCGCATGTCGATGTCGCAACGATCAGATGGCGCTGGCGCAA 842
Qy 841 TGCGCGCATTTACCGAGTCCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 843 TGCGCGCATTTACCGAGTCCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACG 902
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGATTTTC 960
Db 903 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGATTTTC 962
Qy 961 GCCTGTGGGGCAAAACGAGGTGGACCGCTTGCTGCAACTCTCTCAGGGCGAGCGGTGA 1020
Db 963 GCCTGTGGGGCAAAACGAGGTGGACCGCTTGCTGCAACTCTCTCAGGGCGAGCGGTGA 1022
Qy 1021 AGGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGGCGCCCAATA 1080
Db 1023 AGGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGGCGCCCAATA 1082
Qy 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGACAGCTGGCAACGACAGTTT 1140
Db 1083 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGACAGCTGGCAACGACAGTTT 1142
Qy 1141 CCCGACTGGAAGCGGGGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1143 CCCGACTGGAAGCGGGGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1202
Qy 1201 GCAC 1204
Db 1203 GCAC 1206

RESULT 9
US-08-487-283A-18/c
; Sequence 18, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.

APPLICANT: Matis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Suinto, Stephen P.
APPLICANT: Thomas, Thomas C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: 25 Science Park (Alexion)
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Cetrus 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)776-1790
TELEFAX: (203)772-3655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: pET Trc SOS/NI
DESCRIPTION: prokaryotic expression vector
US-08-487-283A-18

Query Match 36.3%; Score 1199.2; DB 3; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGGACACCATCAATGGTGCAAAACCTTTCGGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 4810 CGGACACCATCAATGGCGCAAAACCTTTCGGCGGTATGGCATGATAGCGCCCGGAAGAGA 4751
QY 61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTATACGATTCGAGAGTATGCGG 120
DB 4750 GTCAATTACAGGTGGTGAATGTGAACACCAAGTATACGATTCGAGAGTATGCGG 4691
QY 121 GTGTCTCTTATCAGACCGTTTCGGCGGTGTGAACACCAAGTATGCGGCGGCAAA 180
DB 4690 GTGTCTCTTATCAGACCGTTTCGGCGGTGTGAACACCAAGTATGCGGCGGCAAA 4631
QY 181 CGCGGGAAGTGAAGCGGCGATGGCGGAGTGAATTAATTCACACCGCGTGGCAC 240
DB 4630 CGCGGGAAGTGAAGCGGCGATGGCGGAGTGAATTAATTCACACCGCGTGGCAC 4571
QY 241 AACAACTGGCGGCAACAGTGTGTCTGATTTGGCGTTCCACCTCCAGTCTGGCCCTGC 300
DB 4570 AACAACTGGCGGCAACAGTGTGTCTGATTTGGCGTTCCACCTCCAGTCTGGCCCTGC 4511
QY 301 ACGCGCCGTGCGAAATTTGTCGGCGGCGAATTAATTCGCGCGCGATCAATCTGGGTGCCAGCG 360

Db 4510 ACGGCGCGTCGAAATTTGTCGGCGGATTAATCTCGCGCGCATCAATCTGGTGCCAGCG 4451
QY 361 TGGTGGTGTGATGGTGAACGAAGCGCGTGAAGCGCTTAAAGCGCGGTGCAACAATC 420
Db 4450 TGGTGGTGTGATGGTGAACGAAGCGCGTGAAGCGCTTAAAGCGCGGTGCAACAATC 4391
QY 421 TTCTCGCGCAACGGGTGAGTGGGCTGATCAATTAATCTCCGTGGATCAACGAGTGC 480
Db 4390 TTCTCGCGCAACGGGTGAGTGGGCTGATCAATTAATCTCCGTGGATCAACGAGTGC 4331
QY 481 TTCTGTGTGAAGCTGCCTGCACATAATGTTCCGGCGTATTTCTTCTGATCTCTGACCA 540
Db 4330 TTCTGTGTGAAGCTGCCTGCACATAATGTTCCGGCGTATTTCTTCTGATCTCTGACCA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGGACTGGCGGTGGAGCATC 600
Db 4270 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGGACTGGCGGTGGAGCATC 4211
QY 601 TGGTGGCATTTGGGTCAACGAAATCGCGCTGTTAGCGGCGCCCATTAAGTTCTGTCTCGG 660
Db 4210 TGGTGGCATTTGGGTCAACGAAATCGCGCTGTTAGCGGCGCCCATTAAGTTCTGTCTCGG 4151
QY 661 CGCGTCTCGCTTGGCTGGGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4150 CGCGTCTCGCTTGGCTGGGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCCGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 4090 CGGAACGGGAAGCGGACTGGAGTGCATGTCCGTTTTTCAACAAACCATGCAAAATGCTGA 4031
QY 781 ATAGGGCATGTTCCCACTCGGATGCTGTTGCCAAGCATCAGATGGCGGTGGCGGCA 840
Db 4030 ATAGGGCATGTTCCCACTCGGATGCTGTTGCCAAGCATCAGATGGCGGTGGCGGCA 3971
QY 841 TGGCGCCATTTACCGAGTCCGGGCTGCGGCTTGGTGGCGATATCTCGGTAGTGGATAG 900
Db 3970 TGGCGCCATTTACCGAGTCCGGGCTGCGGCTTGGTGGCGATATCTCGGTAGTGGATAG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATATCCGCGCTTAAACCAACCATCAACAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATATATCCGCGCTTAAACCAACCATCAACAGGATTTTC 3851
QY 961 GCCTGCTGGGCAACACGAGCGTGGACCGCTTGTGCAAACTCTCTCAGGCGCAGCGGTGA 1020
Db 3850 GCCTGCTGGGCAACACGAGCGTGGACCGCTTGTGCAAACTCTCTCAGGCGCAGCGGTGA 3791
QY 1021 AGGCAATCAGCTGTTGCCCGCTCTCACCTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 3790 AGGCAATCAGCTGTTGCCCGCTCTCACCTGGTGAAGAAACCAACCCCTGGCGCCCAATA 3731
QY 1081 CGCAACCGCCTCTCCCGCGCTTGGCGGATTCATTAATGCGCTGGCAGCAGAGTTT 1140
Db 3730 CGCAACCGCCTCTCCCGCGCTTGGCGGATTCATTAATGCGCTGGCAGCAGAGTTT 3671
QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGAGTGTAGTCTCATTTAG 1200
Db 3670 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGAGTGTAGTCTCATTTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607
RESULT 10
PCT-US96-05611A-21/c
; Sequence 21, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Leonardo, Michael J.
; APPLICANT: McFarland, Henry P.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pBT-cer
US-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGC	AAACCTTTCGGGTATGGCATGATAGCGCCGGAAGAGA	60
DB	4622	CGGACACCATCGAATGGCGCA	AAACCTTTCGGGTATGGCATGATAGCGCCGGAAGAGA	4563
QY	61	GTCAATTACAGGTGGTGAAT	GTGAACACGATGAACGTTATACGATGTGCGAGAGTATGCCG	120
DB	4562	GTCAATTACAGGTGGTGAAT	GTGAACACGATGAACGTTATACGATGTGCGAGAGTATGCCG	4503
QY	121	GTGTCCTTATCAGACCGGTT	CCCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180
DB	4502	GTGTCCTTATCAGACCGGTT	CCCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	4443
QY	181	CGCGGGAAGAGTGAAGCGG	CGATGGCGGAGCTGAATTACATTCGCCAACCGGTGGCAC	240
DB	4442	CGCGGGAAGAGTGAAGCGG	CGATGGCGGAGCTGAATTACATTCGCCAACCGGTGGCAC	4383
QY	241	AACAACTCGCGGCAAAAC	AGTCTGTTGCTGATTGGCGTTGGCCACCTCCAGTCTGGCCCTGC	300
DB	4382	AACAACTCGCGGCAAAAC	AGTCTGTTGCTGATTGGCGTTGGCCACCTCCAGTCTGGCCCTGC	4323
QY	301	AGCGCCGTTCGCAAAATGT	CGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
DB	4322	AGCGCCGTTCGCAAAATGT	CGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	4263
QY	361	TGGTGGTTCGATGTAGAAC	GAAGCGGCGTGAAGCGGCTGTAAAGCGGCGGTGCACAATC	420
DB	4262	TGGTGGTTCGATGTAGAAC	GAAGCGGCGTGAAGCGGCTGTAAAGCGGCGGTGCACAATC	4203
QY	421	TTCTCGCGCAACGCGTCAG	TGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA	480
DB	4202	TTCTCGCGCAACGCGTCAG	TGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA	4143
QY	481	TTGCTGTGGAAGCTGCCTG	CACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA	540
DB	4142	TTGCTGTGGAAGCTGCCTG	CACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA	4083
QY	541	CACCCATCAACGATTAATTT	CTCCATGAAGACGGTACCGCATGGCGGTGGAGCATC	600
DB	4082	CACCCATCAACGATTAATTT	CTCCATGAAGACGGTACCGCATGGCGGTGGAGCATC	4023
QY	601	TGGTGCATTTGGTCCAGCA	BAATCGCGCTGTAGCGGCGCCATTAAAGTTCGTCTCGG	660
DB	4022	TGGTGCATTTGGTCCAGCA	BAATCGCGCTGTAGCGGCGCCATTAAAGTTCGTCTCGG	3963
QY	661	CCGCTCTGCGTCTGGCTGG	CTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	3962	CCGCTCTGCGTCTGGCTGG	CTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	3903
QY	721	CGGAAACGGGAACGCGAC	TGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA	780
DB	3902	CGGAAACGGGAACGCGAC	TGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA	3843
QY	781	ATGAGGGCATCGTCCCACT	CGGATGCTGGTTCGCCAAGATGAGTGGCGTGGCGGCA	840
DB	3842	ATGAGGGCATCGTCCCACT	CGGATGCTGGTTCGCCAAGATGAGTGGCGTGGCGGCA	3783

QY	841	TGCGCGCCATTACCGAGTCC	GGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGATACG	900
DB	3782	TGCGCGCCATTACCGAGTCC	GGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGATACG	3723
QY	901	ACGATACCGAAGACAGCTCA	TGTTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC	960
DB	3722	ACGATACCGAAGACAGCTCA	TGTTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC	3663
QY	961	GCCTGCTGGGCAAAACCA	CGCTGACGCTTGTCTCTCAGGGCCAGCGGTGA	1020
DB	3662	GCCTGCTGGGCAAAACCA	CGCTGACGCTTGTCTCTCAGGGCCAGCGGTGA	3603
QY	1021	AGGGCAATCAGCTGTTGCC	CGCTCTCACTGGTGAAGAAAAACCACTTGGGGCCCAATA	1080
DB	3602	AGGGCAATCAGCTGTTGCC	CGCTCTCACTGGTGAAGAAAAACCACTTGGGGCCCAATA	3543
QY	1081	CGCAAAACCGCTCTCTCC	CGCGCTTGGCGGATTCATTAATGCAGCTGCGACAGAGTTT	1140
DB	3542	CGCAAAACCGCTCTCTCC	CGCGCTTGGCGGATTCATTAATGCAGCTGCGACAGAGTTT	3483
QY	1141	CCGACTGGAAGCGGGCAG	GTGAGCGCAACCAATTAATGTAGTTAGTCTCACTCATTAG	1200
DB	3482	CCGACTGGAAGCGGGCAG	GTGAGCGCAACCAATTAATGTAGTTAGTCTCACTCATTAG	3423
QY	1201	GCAC	1204	
DB	3422	GCAC	3419	

RESULT 12

US-08-929-967-1/c
; Sequence 1, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Ruppert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCACAAACCTTTCCGGGTATGGCATGTAGCGCCCGGAAGAGA 60
Db |||||
4810 CGGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGGCATGTAGCGCCCGGAAGAGA 4751
QY 61 GTCAATTCAGGCTGGTGAATGTGAACACAGTAACTGATACGATGTCGAGAGTATGCCG 120
Db |||||
4750 GTCAATTCAGGCTGGTGAATGTGAACACAGTAACTGATACGATGTCGAGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGCTTTCCGGGTGTGAACCCAGGCGCACCGCTTTCTGCGAAAA 180
Db |||||
4690 GTGTCTCTTATCAGACCGCTTTCCGGGTGTGAACCCAGGCGCACCGCTTTCTGCGAAAA 4631
QY 181 CCGGGGAAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db |||||
4630 CCGGGGAAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 4571
QY 241 AACAACTGGCGGCAACAGTGTGCTGATTTGGGTGGCCACCTCCAGTCTGGCCCTGC 300
Db |||||
4570 AACAACTGGCGGCAACAGTGTGCTGATTTGGGTGGCCACCTCCAGTCTGGCCCTGC 4511
QY 301 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 360
Db |||||
4510 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 4451
QY 361 TGGTGGTTCGATGTAGAACGAGCGGCTGAAAGCTGTAAAGCGCGGTGCACAATC 420
Db |||||
4450 TGGTGGTTCGATGTAGAACGAGCGGCTGAAAGCTGTAAAGCGCGGTGCACAATC 4391
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAAGTATCCGCGGCTTCTGATCTCTGACAG 480
Db |||||
4390 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAAGTATCCGCGGCTTCTGATCTCTGACAG 4331
QY 481 TTGCTGTGGAAGCTCCCTGCAATAATGTTCGGGCTTATTTCTTGTGATCTCTGACAGAGA 540
Db |||||
4330 TTGCTGTGGAAGCTCCCTGCAATAATGTTCGGGCTTATTTCTTGTGATCTCTGACAGAGA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAGACGGTACCGGATCGGCGGTGAGCATC 600
Db |||||
4270 CACCCATCAACAGTATTAATTTCTCCCATGAGACGGTACCGGATCGGCGGTGAGCATC 4211
QY 601 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCATTAAGTCTCTCTCGG 660
Db |||||
4210 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCATTAAGTCTCTCTCGG 4151
QY 661 CGCGTCTCGCTCGCTGGCTGGCATTAATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db |||||
4150 CGCGTCTCGCTCGCTGGCTGGCATTAATCTCACTCGCAATCAAAATCAGCCGATAG 4091
QY 721 CCGAACCGGAAAGGCACTGGAGTGCCATGTCCGGTCTTCAACCAACCATGCNAATGCTGA 780
Db |||||
4090 CCGAACCGGAAAGGCACTGGAGTGCCATGTCCGGTCTTCAACCAACCATGCNAATGCTGA 4031
QY 781 ATGAGGGATCGTTCCACATCGCATGTGGTTGCCAAGCATGAGTGGCGCTGGCGCAA 840
Db |||||
4030 ATGAGGGATCGTTCCACATCGCATGTGGTTGCCAAGCATGAGTGGCGCTGGCGCAA 3971
QY 841 TGCAGCATTTACCGAGTCCGGGCTGGCTGGTGGGATATCTCGGTAGTGGATACG 900
Db |||||
3970 TGCAGCATTTACCGAGTCCGGGCTGGCTGGTGGGATATCTCGGTAGTGGATACG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAACACAGATTTTC 960
Db |||||
3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAACACAGATTTTC 3851
QY 961 GCCTGCTGGGGCAACACAGGTGGAACCGTTGTCGAACTCTCTCAGGGCCAGCGGTGA 1020
Db |||||
3850 GCCTGCTGGGGCAACACAGGTGGAACCGTTGTCGAACTCTCTCAGGGCCAGCGGTGA 3791
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCGCTGGCGCCCAATA 1080
Db |||||
3790 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCGCTGGCGCCCAATA 3731
QY 1081 CGCAAAACCGGCTCTCCCGCGCGCTTGGCGGATTCATTAATGCAAGCTGGCACGACAGGTTT 1140

Db 3730 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCAGTCGACGACAGGTTT 3671
QY 1141 CCCGACTGGAAGCGGCGAGTGACGCAACGCAATTAATGTAGTTCAGTCTCATTAG 1200
Db |||||
3670 CCCGACTGGAAGCGGCGAGTGACGCAACGCAATTAATGTAGTTCAGTCTCATTAG 3611
QY 1201 GCAC 1204
Db |||||
3610 GCAC 3607

RESULT 13
US-09-702-705-785/c
; Sequence 785, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCACAAACCTTTCCGGGTATGGCATGTAGCGCCCGGAAGAGA 60
Db |||||
4677 CGGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGGCATGTAGCGCCCGGAAGAGA 4618
QY 61 GTCAATTCAGGCTGGTGAATGTGAACCAAGTAACTGATACGATGTGCAGAGTATGCCG 120
Db |||||
4617 GTCAATTCAGGCTGGTGAATGTGAACCAAGTAACTGATACGATGTGCAGAGTATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGGTGGTGAACCCAGGCGCACCGTTCGCGAAAA 180
Db |||||
4557 GTGTCTCTTATCAGACCGCTTTCCCGGTGGTGAACCCAGGCGCACCGTTCGCGAAAA 4498
QY 181 CCGGGGAAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db |||||
4497 CCGGGGAAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGCAACAGTGTGCTGATTTGGGTGGCCACCTCCAGTCTGGCCCTGC 300
Db |||||
4437 AACAACTGGCGGCAACAGTGTGCTGATTTGGGTGGCCACCTCCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 360
Db |||||
4377 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 4318
QY 361 TGGTGGTTCGATGTAGAACGAGCGGCTGAAAGCTGTAAAGCGCGGTGCACAATC 420
Db |||||
4317 TGGTGGTTCGATGTAGAACGAGCGGCTGAAAGCTGTAAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTCATTAATGCAAGCTGGCACGACAGGTTT 480

Db 4257 TTCTCGCGCAACGGCTCAGTGGCTGATCAATTAACATATCCGCTGGATGACACAGGATGCCA 4198
QY 481 TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db 4197 TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 4138
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCACTGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCACTGGCGGTGGAGCATC 4078
QY 601 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 4077 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 4018
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCACGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCACGCCGATAG 3958
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTCACAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTCACAAACCATGCAAAATGCTGA 3898
QY 781 ATGAGGGCATGTTCCCACTCGGATGTCGCTGTTGCCAAGCATGAGTGGCGCGCAA 840
Db 3897 ATGAGGGCATGTTCCCACTCGGATGTCGCTGTTGCCAAGCATGAGTGGCGCGCAA 3838
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGCGGATATCTCGGTAGTGGGATAG 900
Db 3837 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGCGGATATCTCGGTAGTGGGATAG 3778
QY 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 960
Db 3777 AGGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTGCTGCACTCTCTCAGGCGCACGGCGTGA 1020
Db 3717 GCCTGCTGGGGCAACACAGCGTGGACCGCTGCTGCACTCTCTCAGGCGCACGGCGTGA 3658
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAAACCCCTGGCGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAAACCCCTGGCGCCCAATA 3598
QY 1081 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGACGTGGCACGACAGGTTT 1140
Db 3597 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGACGTGGCACGACAGGTTT 3538
QY 1141 CCGCACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGTCAGTTAGTCACTCATTAG 1200
Db 3537 CCGCACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGTCAGTTAGTCACTCATTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 14
US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736.457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGCCCGCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGGCGCAAAACCTTTTCGCGGTATGGCATGATAGCCCGCGGAAGAGA 4618
QY 61 GTCAATTTCAGGGTGGTGAATGTAACCCAGTAAACGTTATACGATGTCGAGATGATCGG 120
Db 4617 GTCAATTTCAGGGTGGTGAATGTAACCCAGTAAACGTTATACGATGTCGAGATGATCGG 4558
QY 121 GTGTCCTCTTATCAGACCGGTTTCCCGCGTGTGAACCCAGCCAGCCACGCTTTCTGCGAAAA 180
Db 4557 GTGTCCTCTTATCAGACCGGTTTCCCGCGTGTGAACCCAGCCAGCCACGCTTTCTGCGAAAA 4498
QY 181 CGCGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGGTTTGGCCACTTCCAGTCTGGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGGTTTGGCCACTTCCAGTCTGGGCCCTGC 4378
QY 301 ACGGCGCGTCCCAAAATTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 4377 ACGGCGCGTCCCAAAATTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 4318
QY 361 TGGTGGTTCGATGGTAGAAGCAAGCGCGCTCGAAGCCCTGAAAGCGCGGTGCAACAATC 420
Db 4317 TGGTGGTTCGATGGTAGAAGCAAGCGCGCTCGAAGCCCTGAAAGCGCGGTGCAACAATC 4258
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480
Db 4257 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 4198
QY 481 TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db 4197 TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 4138
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 4078
QY 601 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 4077 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 4018
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCACGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCACGCCGATAG 3958
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTCACAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTCACAAACCATGCAAAATGCTGA 3898
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGCGCGCAA 840
Db 3897 ATGAGGGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGCGCGCAA 3838
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGCGGATATCTCGGTAGTGGGATAG 900
Db 3837 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGCGGATATCTCGGTAGTGGGATAG 3778

Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGGATTTTC	960
Db	3777	ACGATACCGAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGGATTTTC	3718
Qy	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	3717	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	3658
Qy	1021	AGGCAATACAGCTGTTGTCGGCTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA	1080
Db	3657	AGGCAATACAGCTGTTGTCGGCTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA	3598
Qy	1081	CGCAAAACCGGCTCTCTCCCGCGCGTGGCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Db	3597	CGCAAAACCGGCTCTCTCCCGCGCGTGGCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	3538
Qy	1141	CCGCACTGGAAAGCGGGCAGTGACGCCAACGCAATTAATGTAGTTAGCTCACTCAATTAG	1200
Db	3537	CCGCACTGGAAAGCGGGCAGTGACGCCAATTAATGTAGTTAGCTCACTCAATTAG	3478
Qy	1201	GCAC 1204	
Db	3477	GCAC 3474	

```

RESULT 15
US-09-614-124B-785/c
/ Sequence 785, Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedwick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C9
/ CURRENT APPLICATION NUMBER: US/09/614, 124B
/ CURRENT FILING DATE: 2001-07-11
/ NUMBER OF SEQ ID NOS: 1668
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 785
/ LENGTH: 5502
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-614-124B-785

```

Query Match	36.3%	Score 1199.2	DB 4	Length 5502
Best Local Similarity	99.8%	Prod. No. 0		
Matches 1201	Conservative 0	Mismatches 3	Indels 0	Gaps 0
QY	1	CCGACACATCGAATGGTGC	AAACCTTTCCGGGTATG	CATGATAGCCCGGACGAGA 60
Db	4677	CGGACACATCGAATGGCG	CAAAACCTTTCCGGGTATG	CATGATAGCCCGGACGAGA 4618
QY	61	GTCAATTCAGGTGGTGAAT	TGTGAACACAGTAAACG	TATATACGATGTCGACGATATCCG 120
Db	4617	GTCAATTCAGGTGGTGAAT	TGTGAACACAGTAAACG	TATATACGATGTCGACGATATCCG 4558
QY	121	GTGTCTCTTATCAGACCG	TTTCCCGCGTGGTGAAC	CCAGCCAGCCACGCTTTCTCGGAAA 180
Db	4557	GTGTCTCTTATCAGACCG	TTTCCCGCGTGGTGAAC	CCAGCCAGCCACGCTTTCTCGGAAA 4498
QY	181	CGCGGGAAAAGTGGGAAG	CGGCGATCGCGAGCTCA	TATATATCCCAACCGCGTGCAC 240
Db	4497	CGCGGGAAAAGTGGGAAG	CGGCGATCGGCGAGCTCA	TATATATCCCAACCGCGTGCAC 4438
QY	241	AACAACTGGCGGGCAAC	ACAGTCTGCTGATTTGG	CGGTTGCCACCTTCAGTCTGC
				CGGCTGC 300

Db	4437		AACAACTGGCGGGCAACAGTCGTGTTGCTGATGTGGCGTTGCCACTCTCAAGTCTGGCCCTGC	4378
Qy	301		ACGGCCGTCGCAAAATTGTTCGCGCGATTAATCTCGCGCGATCAACTCGGGTGCACGC	360
Db	4377		ACGGCCGTCGCAAAATTGTTCGCGCGATTAATCTCGCGCGATCAACTCGGGTGCACGC	4318
Qy	361		TGTTGGTTCGATGGTAGAACGAACGCGCGTCGAAGCTGTAAAGCGCGGTGCACAATC	420
Db	4317		TGTTGGTTCGATGGTAGAACGAACGCGCGTCGAAGCTGTAAAGCGCGGTGCACAATC	4258
Qy	421		TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAACATTCGCTCGATGACACAGATGCCA	480
Db	4257		TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAACATTCGCTCGATGACACAGATGCCA	4198
Qy	481		TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCCTTGATGTCTCTGACCAGA	540
Db	4197		TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCCTTGATGTCTCTGACCAGA	4138
Qy	541		CACCCATCAACAGTATTATTTCCTCCATGAGACGGTAGCGACCTGGGCGTGGAGATC	600
Db	4137		CACCCATCAACAGTATTATTTCCTCCATGAGACGGTAGCGACCTGGGCGTGGAGATC	4078
Qy	601		TGTTCCGATTGGTTCACAGCAATCCGCTGTTAGCGGGCCCAATTAAGTTCCTGCTCGG	660
Db	4077		TGTTCCGATTGGTTCACAGCAATCCGCTGTTAGCGGGCCCAATTAAGTTCCTGCTCGG	4018
Qy	661		CGGTCCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	4017		CGGTCCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	3958
Qy	721		CGGAAACGGGAAGCGACTGAGTGCACATGCCGTTTCGCGTTTTCAAACAAACCATGCAATCTGA	780
Db	3957		CGGAAACGGGAAGCGACTGAGTGCACATGCCGTTTTCGCGTTTTCAAACAAACCATGCAATCTGA	3898
Qy	781		ATGAGGCGATCGTTCCCACTCGCATGCTGGTTCGCCAACGATCAGATGGCGCTGGGCGCAA	840
Db	3897		ATGAGGCGATCGTTCCCACTCGCATGCTGGTTCGCCAACGATCAGATGGCGCTGGGCGCAA	3838
Qy	841		TGCGCGCATTCACGAGTCCGGCTCGGGTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	3837		TGCGCGCATTCACGAGTCCGGCTCGGGTGGGATATCTCGGTAGTGGGATACG	3778
Qy	901		ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCAACCATCAACAGGATTTTC	960
Db	3777		ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCAACCATCAACAGGATTTTC	3718
Qy	961		GCCTGCTGGGGCAAAACGAGGTGGACCGCTTGCTGCAACTCTCTAGGGCGAGCGGTGA	1020
Db	3717		GCCTGCTGGGGCAAAACGAGGTGGACCGCTTGCTGCAACTCTCTAGGGCGAGCGGTGA	3658
Qy	1021		AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAACCAACCTCGGGCCCAATA	1080
Db	3657		AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAACCAACCTCGGGCCCAATA	3598
Qy	1081		CGCAAAACCGCTCTCCCGGGCGTTGGCCGATTCATTAAATGAGCTGGCACGACAGGTTT	1140
Db	3597		CGCAAAACCGCTCTCCCGGGCGTTGGCCGATTCATTAAATGAGCTGGCACGACAGGTTT	3538
Qy	1141		CCGACTGGAACGGGCGAGTCGACGCAACGCAATTAATGTCAGTTAGTCTCACTATTAG	1200
Db	3537		CCGACTGGAACGGGCGAGTCGACGCAACGCAATTAATGTCAGTTAGTCTCACTATTAG	3478
Qy	1201		GCAC	1204
Db	3477		GCAC	3474

Search completed: August 3, 2005, 02:32:22
Job time : 356.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 14:24:04 ; Search time 1094.6 Seconds
(without alignments)
17846.836 Million cell updates/sec

Title: US-09-765-555B-15
Perfect score: 3300
Sequence: 1 ccgacacatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3300	100.0	3300	4 AAD11588	Aad11588 Partial s
2	3300	100.0	3300	4 AAD11589	Aad11589 Partial s
3	3226.4	97.8	3300	4 AAD11592	Aad11592 Partial s
4	3215.2	97.4	3300	4 AAD11591	Aad11591 Partial s
5	3207.2	97.2	3300	4 AAD11590	Aad11590 Partial s
6	2705.8	82.0	8101	4 AAH25681	Aah25681 Nucleotid
7	2705.8	82.0	8101	6 ABL53238	Ab153238 Nucleotid
8	2690.4	81.5	7259	12 ADO23608	Ado23608 DNA encod
9	2690.4	81.5	7322	12 ADO23613	Ado23613 DNA encod
10	2690.4	81.5	7352	12 ADO23603	Ado23603 DNA encod
11	2690.4	81.5	7370	12 ADO23598	Ado23598 DNA encod
12	2690.4	81.5	7370	12 ADO23639	Ado23639 DNA encod
13	2690.4	81.5	7370	12 ADO23649	Ado23649 DNA encod
14	2690.4	81.5	7370	12 ADO23644	Ado23644 DNA encod
15	2690.4	81.5	7403	12 ADO23594	Ado23594 DNA encod
16	2690.4	81.5	7442	12 ADO23590	Ado23590 DNA encod
17	2690.4	81.5	7478	12 ADO23584	Ado23584 DNA encod
18	2688.8	81.5	7112	12 ADO23618	Ado23618 DNA encod
19	2636.8	79.9	6806	6 ABL49925	Ab149925 Maltose b
20	2605.4	79.0	7553	12 ADO23588	Ado23588 DNA encod

21	1609.4	48.8	5558	12 ADL90419	Adl90419 Clostridi
22	1474	44.7	4700	9 ACF06053	Act06053 Vector pJ
23	1472.4	44.6	4700	12 ADL18670	Adl18670 Vector pJ
24	1462.4	44.3	4920	10 ADH73599	Adh73599 Plasmid p
25	1462.4	44.3	4935	10 ADH73598	Adh73598 Plasmid p
26	1462.4	44.3	4945	10 ADH73600	Adh73600 Plasmid p
27	1462.4	44.3	4951	10 ADH73601	Adh73601 Plasmid p
28	1418.2	43.0	5903	8 ABZ23939	Abz23939 Nucleotid
29	1417.6	43.0	1511	6 ABQ73200	Abq73200 E. coli t
C 30	1417.6	43.0	5926	2 AAV32977	Aav32977 Tn7 donor
C 31	1417.6	43.0	5926	6 AAD45059	Aad45059 Transposo
C 32	1417.6	43.0	5926	12 ADG46817	Adg46817 Donor pla
33	1362	41.3	5201	12 ADL72229	Adl72229 DNA seque
34	1362	41.3	5201	12 ADL72228	Adl72228 DNA seque
C 35	1238.4	37.5	5024	9 ACF06051	Act06051 Plasmid p
C 36	1238.4	37.5	5024	12 ADL18668	Adl18668 Plasmid p
37	1237.2	37.5	1922	6 ABQ73089	Abq73089 MBP and z
38	1237.2	37.5	1922	10 ADF83622	Adf83622 MBP-human
39	1204	36.5	4557	2 AAT90491	Aat90491 Vector pl
40	1204	36.5	4969	12 ADO33522	Ado33522 Expressio
41	1204	36.5	5006	10 ADE24112	Ade24112 plasmid p
42	1204	36.5	5064	12 ADO33532	Ado33532 Vector pG
43	1204	36.5	5082	12 ADO33531	Ado33531 Vector pG
44	1204	36.5	6259	6 AAD29720	Aad29720 Plasmid p
45	1204	36.5	6823	3 AAC55459	Aac55459 Destinati

ALIGNMENTS

RESULT 1

AAD11588

ID AAD11588 standard; DNA; 3300 BP.

XX

AC AAD11588;

XX

DT 24-SEP-2001 (first entry)

XX

DE Partial sequence of pMal-m1 and ZFPm1 DNA.

XX

KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;

XX modulation; plant technology; agriculture; ds.

XX

OS Unidentified.

XX

PH Key Location/Qualifiers

FT CDS

FT 2719..3270

FT

FT /tag= a

FT /product= "ZFPm1 protein"

FT /transl_except= (pos:2785..2787, aa:Asp)

FT /transl_except= (pos:2788..2790, aa:Pro)

FT /transl_except= (pos:2791..2793, aa:Gly)

FT /transl_except= (pos:2872..2874, aa:Arg)

FT /transl_except= (pos:2875..2877, aa:Ala)

FT /transl_except= (pos:2878..2880, aa:His)

FT /transl_except= (pos:2884..2886, aa:Glu)

FT /transl_except= (pos:2953..2955, aa:Gln)

FT /transl_except= (pos:2959..2961, aa:Ser)

FT /transl_except= (pos:3121..3123, aa:Arg)

FT /transl_except= (pos:3124..3126, aa:Ser)

FT /transl_except= (pos:3127..3129, aa:Asp)

FT /transl_except= (pos:3130..3132, aa:Asn)

FT /transl_except= (pos:3136..3138, aa:Val)

FT /transl_except= (pos:3139..3141, aa:Arg)

FT /transl_except= (pos:3205..3207, aa:Gln)

FT /transl_except= (pos:3208..3210, aa:Ala)

FT /transl_except= (pos:3211..3213, aa:Gly)

FT /transl_except= (pos:3214..3216, aa:His)

FT /transl_except= (pos:3220..3222, aa:Ala)

FT /transl_except= (pos:3223..3225, aa:Ser)

FT /note= "CDS does not include start and stop codon"

FT /partial

FT primer_bind

FT 2740..2790


```

FT primer_bind /tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /tag= c
FT primer_bind /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /tag= d
FT primer_bind /bound_moiety= "F2-b primer"
FT 2867..2940
FT /tag= e
FT primer_bind /bound_moiety= "F2-f primer"
FT 2916..2973
FT /tag= f
FT primer_bind /bound_moiety= "F3-b1 primer"
FT 2953..3021
FT /tag= g
FT primer_bind /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT /tag= h
FT primer_bind /bound_moiety= "F4-f2 primer"
FT 3022..3102
FT /tag= i
FT primer_bind /bound_moiety= "F4-f1 primer"
FT 3076..3141
FT /tag= j
FT primer_bind /bound_moiety= "F5-b primer"
FT 3119..3192
FT /tag= k
FT primer_bind /bound_moiety= "F5-f primer"
FT 3168..3225
FT /tag= l
FT primer_bind /bound_moiety= "F6-b1 primer"
FT 3205..3273
FT /tag= m
FT primer_bind /bound_moiety= "F6-b2 primer"

```

WO200152620-A2.

26-JUL-2001.

19-JAN-2001; 2001WO-US001817.

21-JAN-2000; 2000US-0177468P.

21-JUL-2000; 2000US-00620897.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA AGRIC DISCOVERY INC.

Barbas CF, Stege JT, Guan X, Dalmia B;

WPI; 2001-465325/50.

P-PSDB; AAE06000.

New zinc finger proteins, useful for modulating or regulating gene expression and metabolic pathways in plants, e.g. for treating in the plant cells a disorder that is associated with abnormal expression of the target gene.

Example 4; Page 138-139; 156pp; English.

The patent discloses methods and compositions to modulate the expression of a target gene in plant cells. The method involves providing plant cells with a zinc finger protein (ZFP) which is capable of specifically binding to a target nucleotide sequence or its complementary strand within a target gene and allowing the ZFP binding to the target nucleotide sequence, where the expression of the target gene in the plant cells is modulated. The ZFP and fusions of the ZFP proteins are useful for modulating or regulating gene expression and metabolic pathways in plants. The ZFP, fusion proteins and methods are useful in plant and agricultural technology. The method is useful particularly for treating a disorder in the plant cells, where the disorder is associated with abnormal expression of the target gene. The present DNA sequence is the partial sequence of pMal-m1 and ZFPm1 DNA

```

XX SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
Query Match 100.0%; Score 3300; DB 4; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACATCGAATGTTGCAAAACCTTTTCGGGTATGGCATGATACGCCCGGAAGAGA 60
DB 1 CCGACACATCGAATGTTGCAAAACCTTTTCGGGTATGGCATGATACGCCCGGAAGAGA 60
QY 61 GTCATTTTCAGGTGCTGAATGTGAACACAGTAACTTATACGATGTCGCGAGATATCCG 120
DB 61 GTCATTTTCAGGTGCTGAATGTGAACACAGTAACTTATACGATGTCGCGAGATATCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCGAGCCAGCTTCTGCGAAAA 180
DB 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCGAGCCAGCTTCTGCGAAAA 180
QY 181 CCGGGGAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTCCTCCACCGCGTGGCAC 240
DB 181 CCGGGGAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTCCTCCACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGCGCGTTGCCACTCCAGTCTGSCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGCGCGTTGCCACTCCAGTCTGSCCTGC 300
QY 301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGCGGTGCACATC 420
DB 361 TGGTGTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGCGGTGCACATC 420
QY 421 TTCTCGCGCAACCGGTCACTGGGCTGATCAATTAATCTATCGCTGGATGACACGAGTCCCA 480
DB 421 TTCTCGCGCAACCGGTCACTGGGCTGATCAATTAATCTATCGCTGGATGACACGAGTCCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTTCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTTCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGGGTGGAGCATC 600
DB 541 CACCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGGGTGGAGCATC 600
QY 601 TGGTTCGATTTGGTTCACAGCAAAATCGCGTGTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTTCGATTTGGTTCACAGCAAAATCGCGTGTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CCGGTCTGGCTTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
DB 661 CCGGTCTGGCTTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
QY 721 CCGAAGCGGAAGCGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATCTGA 780
DB 721 CCGAAGCGGAAGCGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATCTGA 780
QY 781 ATGAGGCGCATCGTTCCCACTGGCATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
DB 781 ATGAGGCGCATCGTTCCCACTGGCATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGCGCATTAACGAGTCCGGGTGCGGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGCGCATTAACGAGTCCGGGTGCGGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 AGCATACCGAAGACAGCTCATGTTATATCCCGCGTTTAACACCATCAAAACAGGATTTTC 960
DB 901 AGCATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACACCATCAAAACAGGATTTTC 960
QY 961 GCGTGTGGGGCAACACAGCTGGACCGGTGCTGCTCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 961 GCGTGTGGGGCAACACAGCTGGACCGGTGCTGCTCAACTCTCTCAGGGCCAGCGGTGA 1020

```

Db 961 GCCTGCTGGGCAACACGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAAAACCACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAAAACCACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTGGCCGATTCAATGAGCTGGCGACACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTGGCCGATTCAATGAGCTGGCGACACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTCAGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTCAGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Qy 1201 GCACAATCTCATGTTTGACGCTTATCATGCACTGACGCTGCAACGATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACGCTTATCATGCACTGACGCTGCAACGATGCTTCTGGCG 1260
Qy 1261 TCAGGCAGCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTGG 1320
Db 1261 TCAGGCAGCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTGG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATTCTGAATGAGCTGTGCAATTAATCAATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAATATTCTGAATGAGCTGTGCAATTAATCAATCGGCTCGTATAATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTCACACAGAAACAGCCAGTCGGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTCACACAGAAACAGCCAGTCGGTTTAGTGTTCACGA 1500
Qy 1501 GCACCTTCACCAACAGGACCATAGATTATGAAGAACTGAAGAAAGTTAACTGGTAATCTGG 1560
Db 1501 GCACCTTCACCAACAGGACCATAGATTATGAAGAACTGAAGAAAGTTAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGNAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGNAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACGACCGCTTGGTGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACGACCGCTTGGTGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGAACAACGCTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGAACAACGCTTCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCCGTAGCTTACAACGGCAAGCTGATTCCTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCCGTAGCTTACAACGGCAAGCTGATTCCTACCGGATCGCTGTT 1860
Qy 1861 GAAAGCTTATCGTGAATTTAACAAGATCTGTCGCGAACCAGCAACCTGGGAA 1920
Db 1861 GAAAGCTTATCGTGAATTTAACAAGATCTGTCGCGAACCAGCAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGCGTGAATGCTGCTGACGCGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCGTGAATGCTGCTGACGCGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAAGACGTGGGCGTGGATAACCGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAAGACGTGGGCGTGGATAACCGCTGGCGGAAAGCG 2100

Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAAGCTGCCCTTTAATAAAGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCCCTTTAATAAAGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACACGAGCAAAAGTGAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACACGAGCAAAAGTGAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTTCAACCATCCAAACCGTTGCTGGCGTCTGAGCGCAGGTTAATAACGCCGCCAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTGCTGGCGTCTGAGCGCAGGTTAATAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAATATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAATATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTAATAAAGACAAACCGCTGGTGCCTGAGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAAGACAAACCGCTGGTGCCTGAGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCAAAAGATCCACGTTATTTGCCGCCACCATGGAACCGCCAGAAAGGTTGAATCATG 2520
Db 2461 TTGGCAAAAGATCCACGTTATTTGCCGCCACCATGGAACCGCCAGAAAGGTTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGTCGCTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGTCGCTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCTCTCTGTCGCGCCAGCGGCTTCGAGCCCGGGAAGGAGCCCTATGCTGT 2760
Db 2701 GGATCCTCTCTCTGTCGCGCCAGCGGCTTCGAGCCCGGGAAGGAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGTAAGTCTCTCTCAGAGCTCTCACCTGGTGGCCACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTCTCTCAGAGCTCTCACCTGGTGGCCACAGCGTACCCAC 2820
Qy 2821 ACGGTTGAAAAACCGTATAAATGCCAGATGGCGCAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTTGAAAAACCGTATAAATGCCAGATGGCGCAATCTTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTCCGCATCAACGCACTCATCTGGGGAAGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGGTCCGCATCAACGCACTCATCTGGGGAAGCCATACAAATGTCCAGATGTGGC 2940
Qy 2941 AAGTCTTTCTCTCGGTCTGCAAAATCTCGTCCGGCACCAACGTAATCTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGCAAAATCTCGTCCGGCACCAACGTAATCTCACACCGGGGAGAG 3000
Qy 3001 CCCTATGTTGTCCGGAATGTTAAGTCTTTCAGCCGAGCGATAAACCCTGGTGGGCCAC 3060
Db 3001 CCCTATGTTGTCCGGAATGTTAAGTCTTTCAGCCGAGCGATAAACCCTGGTGGGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
Qy 3121 CAGGCGGCGCACTGGCCAGCCATCAACGCACTCTACTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGGCGGCGCACTGGCCAGCCATCAACGCACTCTACTGCGGAGAGCCATACAAATGT 3180

QY 3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240
 Db |||||
 3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240
 QY 3241 ACCGGTAAATAAAGTGTGGCCAGGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT 3300
 Db |||||
 3241 ACCGGTAAATAAAGTGTGGCCAGGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT 3300

RESULT 2

AAD11589

ID AAD11589 standard; DNA; 3300 BP.

XX AAD11589;

AC AAD11589;

XX AAD11589;

DT 24-SEP-2001 (first entry)

XX 24-SEP-2001 (first entry)

DE Partial sequence of pMal-m2 and ZFPm2 DNA.

XX Partial sequence of pMal-m2 and ZFPm2 DNA.

KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;

KW modulation; plant technology; agriculture; ds.

XX Unidentified.

OS Unidentified.

XX Unidentified.

FH Key

FT Location/Qualifiers

FT 2719..3270

FT /tag= a

FT /product= "ZFPm2 protein"

FT /note= "CDS does not include start and stop codon"

FT /partial

FT primer_bind

FT 2740..2790

FT /tag= b

FT /bound_moiety= "F1-f2 primer"

FT primer_bind

FT 2770..2850

FT /tag= c

FT /bound_moiety= "F1-f1 primer"

FT primer_bind

FT 2824..2889

FT /tag= d

FT /bound_moiety= "F2-b primer"

FT primer_bind

FT 2867..2940

FT /tag= e

FT /bound_moiety= "F2-f primer"

FT primer_bind

FT 2916..2973

FT /tag= f

FT /bound_moiety= "F3-b1 primer"

FT primer_bind

FT 2953..3021

FT /tag= g

FT /bound_moiety= "F3-b2 primer"

FT primer_bind

FT 2992..3042

FT /tag= h

FT /bound_moiety= "F4-f2 primer"

FT primer_bind

FT 3022..3102

FT /tag= i

FT /bound_moiety= "F4-f1 primer"

FT primer_bind

FT 3076..3141

FT /tag= j

FT /bound_moiety= "F5-b primer"

FT primer_bind

FT 3119..3192

FT /tag= k

FT /bound_moiety= "F5-f primer"

FT primer_bind

FT 3168..3225

FT /tag= l

FT /bound_moiety= "F6-b1 primer"

FT primer_bind

FT 3205..3273

FT /tag= m

FT /bound_moiety= "F6-b2 primer"

FT primer_bind

PR

PR

XX

XX

PA

PA

XX

XX

PI

XX

XX

DR

DR

XX

XX

PT

PT

PT

PT

XX

XX

FS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Query Match

Best Local Similarity 100.0%; Score 3300; DB 4; Length 3300;

Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60

1 CCACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60

61 GTCAATTTCAGGGTGGTGAATGTGAACACAGTACGTTATACGATGTCGACAGTATGCCG 120

61 GTCAATTTCAGGGTGGTGAATGTGAACACAGTACGTTATACGATGTCGACAGTATGCCG 120

121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCGGACCGTTCCTTCGCAAAA 180

121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCGGACCGTTCCTTCGCAAAA 180

181 CGCGGCAAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240

181 CGCGGCAAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240

241 AACACTGGCGGCAAAACAGTCTGTTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

241 AACACTGGCGGCAAAACAGTCTGTTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

301 ACGGCGGTCGCAAAATTTGCGGGCGAATTAATCTCGGCCCGATCAACTGGGTGCGAGCG 360

301 ACGGCGGTCGCAAAATTTGCGGGCGAATTAATCTCGGCCCGATCAACTGGGTGCGAGCG 360

361 TGGTGTGTGTCGATGGTAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCAAAATC 420

361 TGGTGTGTGTCGATGGTAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCAAAATC 420

421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCCCGTGGATGACAGGATGCCA 480

421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCCCGTGGATGACAGGATGCCA 480

481 TTGCTGTGGAAGCTGCTGACTTAATGTTCCGGCGTTATTTCTTGATGTCTTGACCA 540

481 TTGCTGTGGAAGCTGCTGACTTAATGTTCCGGCGTTATTTCTTGATGTCTTGACCA 540

QY 3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240

Db |||||

3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240

QY 3241 ACCGGTAAATAAAGTGTGGCCAGGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT 3300

Db |||||

3241 ACCGGTAAATAAAGTGTGGCCAGGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT 3300

QY 3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240

Db |||||

3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240

QY 3241 ACCGGTAAATAAAGTGTGGCCAGGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT 3300

Db |||||

3241 ACCGGTAAATAAAGTGTGGCCAGGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT 3300

QY 3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240

Db |||||

3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240

QY 3241 ACCGGTAAATAAAGTGTGGCCAGGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT 3300

Db |||||

3241 ACCGGTAAATAAAGTGTGGCCAGGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT 3300

QY	541	CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
QY	601	TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG	660
DB	601	TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG	660
QY	661	CGCGTCTCGCTGCGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTCGCTGCGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGCACTGGAGTGCATTCGCGTGTTCACAAACCAATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGCGCACTGGAGTGCATTCGCGTGTTCACAAACCAATGCAAAATGCTGA	780
QY	781	ATGAGGGCATGTTCCCACTCGATGCTGTGTCACACGATCAGATGCGCTGGCGGCAA	840
DB	781	ATGAGGGCATGTTCCCACTCGATGCTGTGTCACACGATCAGATGCGCTGGCGGCAA	840
QY	841	TGCGGCCATTACCGAGTCCGGGCTGCGCTGGTGGCGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGGCCATTACCGAGTCCGGGCTGCGCTGGTGGCGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC	960
DB	901	ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCCACGGCGTGA	1020
DB	961	GCCTGCTGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCCACGGCGTGA	1020
QY	1021	AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA	1080
DB	1021	AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCGTGGCGACGACAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCGTGGCGACGACAGGTTT	1140
QY	1141	CCCGACTGGAAGCGGCAGTGGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
DB	1141	CCCGACTGGAAGCGGCAGTGGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGCAGCCATCGGAAGCTGGTATGGCTGTGCAGGTTCGTAATCACTGCGATTAATTCG	1320
DB	1261	TCAGGCAGCCATCGGAAGCTGGTATGGCTGTGCAGGTTCGTAATCACTGCGATTAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACTCCGCTTCTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380
DB	1321	TGTCGCTCAAGCGCACTCCGCTTCTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380
QY	1381	CTGGCAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
DB	1381	CTGGCAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGTAGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGTAGTGTTCACGA	1500
QY	1501	GCACCTTACCACACAGGACCATAGATTATGAAAACTGAAGAAAGGTAAACTGTGTAATCTGG	1560
DB	1501	GCACCTTACCACACAGGACCATAGATTATGAAAACTGAAGAAAGGTAAACTGTGTAATCTGG	1560
QY	1561	ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAAGTACGGAATTCAGAAAGAT	1620
DB	1561	ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAAGTACGGAATTCAGAAAGAT	1620

QY	1621	ACCGGAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAGAGAAATTTCCACAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAGAGAAATTTCCACAGGTT	1680
QY	1681	CGGCAACTGGCGATGGCCCTGCATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
DB	1681	CGGCAACTGGCGATGGCCCTGCATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGAACAAGCTGAT	1800
DB	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGAACAAGCTGAT	1800
QY	1801	CCGTTTACCTGGGATGGCTAGCTTAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
DB	1801	CCGTTTACCTGGGATGGCTAGCTTAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
QY	1861	GAACGCTTATCGCTGATTATTAACAAGATCTGCTGCGAAACCGCCCAAAACCTGGGAA	1920
DB	1861	GAACGCTTATCGCTGATTATTAACAAGATCTGCTGCGAAACCGCCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGCTGACGGGGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGCTGACGGGGGTTATCGGTTCAAG	2040
QY	2041	TATGAAACCGCAAGTACGACATTAAGAGCTGGCGGTGGATTAACGCTGGCGCGAAACGG	2100
DB	2041	TATGAAACCGCAAGTACGACATTAAGAGCTGGCGGTGGATTAACGCTGGCGCGAAACGG	2100
QY	2101	GGTCTGACCTTCTCGTTGACCTGATTAAAAAACAAACATGAAATGCAGACACCGATTAC	2160
DB	2101	GGTCTGACCTTCTCGTTGACCTGATTAAAAAACAAACATGAAATGCAGACACCGATTAC	2160
QY	2161	TCCATCGCAAGAGCTGCTTTTAAAGCGCAACACAGCGATGACCATCAACGGCCCGTGG	2220
DB	2161	TCCATCGCAAGAGCTGCTTTTAAAGCGCAACACAGCGATGACCATCAACGGCCCGTGG	2220
QY	2221	GCATGCTCAACATCGACACCGCAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC	2280
DB	2221	GCATGCTCAACATCGACACCGCAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC	2280
QY	2281	AAGGTCACCAACCATCAACCGTTCTGGCGTCTGAGCGCAGGTATTAACCGCCCACT	2340
DB	2281	AAGGTCACCAACCATCAACCGTTCTGGCGTCTGAGCGCAGGTATTAACCGCCCACT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
DB	2341	CCGAACAAAGAGCTGGCAAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
QY	2401	GAAGCGTTTAAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	2460
DB	2401	GAAGCGTTTAAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	2460
QY	2461	TTGGCGAAAGATCCACGTATTTGCGGCCACCATGGAACAAACCGCCAGAAAGTGAATCATG	2520
DB	2461	TTGGCGAAAGATCCACGTATTTGCGGCCACCATGGAACAAACCGCCAGAAAGTGAATCATG	2520
QY	2521	CCGAACATCCCGAGATGTCGCTTTTCTGGTATGCGGTGGTGGTGGTGGTGGTGGTGGTGG	2580
DB	2521	CCGAACATCCCGAGATGTCGCTTTTCTGGTATGCGGTGGTGGTGGTGGTGGTGGTGGTGG	2580
QY	2581	GCAGCGGTCTGACGTGATGAGCCCTGGAAGACGCGACACTTAATTCGAGCTCG	2640
DB	2581	GCAGCGGTCTGACGTGATGAGCCCTGGAAGACGCGACACTTAATTCGAGCTCG	2640
QY	2641	AACAACAAACAAATAACAACTTCGCGGATCGAGGGAAGGATTTTCAAAATTC	2700
DB	2641	AACAACAAACAAATAACAACTTCGCGGATCGAGGGAAGGATTTTCAAAATTC	2700
QY	2701	GGATCCTCTTCTGTTGGCCCAAGCGCGCCCTTCGAGACCCCGGGGAGAACCCCTATGCTGT	2760

```
Db 2701 GGATCCTCTCTCTGTGTGCGCCAGCGCGCCCTCGAGCCGCGGAGAACCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTCTCTCAGAGCTCTACCTGGTGGCCACACAGGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTCTCTCAGAGCTCTACCTGGTGGCCACACAGGTACCCAC 2820
QY 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCATCAACGCACCTATATCTGGCGAGAGCCATCAAAATGTCAGATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACCTATATCTGGCGAGAGCCATCAAAATGTCAGATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCAACGTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCAACGTACTCACACCGGGGAGAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTTCAGCCGCGAGGATAACCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTTCAGCCGCGAGGATAACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTAGC 3120
QY 3121 CAGGCGGCGCACCTGGCGAGCCATCAACGCACCTATCTGCGGAGAGCCATCAAAATGT 3180
Db 3121 CAGGCGGCGCACCTGGCGAGCCATCAACGCACCTATCTGCGGAGAGCCATCAAAATGT 3180
QY 3181 CCAGAATGTGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGCGCACCAACGTACTCAC 3240
Db 3181 CCAGAATGTGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGCGCACCAACGTACTCAC 3240
QY 3241 ACCGGTAAAAAATAGTGGCGAGCGCGCCAGTACCGGTACGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCGAGCGCGCCAGTACCGGTACGAGTTCGGGACTACGCT 3300

RESULT 3
ID AAD11592 standard; DNA; 3300 BP.
XX AC AAD11592;
XX DT 24-SEP-2001 (first entry)
XX DE Partial sequence of pMal- $\alpha$ 3 and ZFP $\alpha$ 3 DNA.
XX KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
XX KW modulation; plant technology; agriculture;  $\alpha$ 3; APETALA3; ds.
XX OS Unidentified.
XX PH Key
XX FT Location/Qualifiers
XX CDS 2719..3320
XX /tag= a
XX /product= "ZFP $\alpha$ 3 protein"
XX /transl_except= (pos:3124..3126, aa:Ser)
XX /transl_except= (pos:3127..3129, aa:Ser)
XX /transl_except= (pos:3136..3138, aa:Val)
XX /note= "CDS does not include start and stop codon"
XX /partial
XX primer_bind 2740..2790
XX /tag= b
XX /bound_moiety= "F1-f2 primer"
XX primer_bind 2770..2850
XX /tag= c
XX /bound_moiety= "F1-f1 primer"
XX primer_bind 2824..2889
XX /tag= d
XX /bound_moiety= "F2-b primer"
```

```
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3188..3225
FT /tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /tag= m
FT /bound_moiety= "F6-b2 primer"
XX W0200152620-A2.
XX 26-JUL-2001.
XX 19-JAN-2001; 2001WO-US001817.
XX 21-JAN-2000; 2000US-0177468P.
XX 21-JUL-2000; 2000US-00620897.
XX (SCRI ) SCRIPPS RES INST.
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX Barbas CF, Stege JT, Guan X, Dalmia B;
XX WPI; 2001-465325/50.
XX P-PSDB; AAE06005.
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX Example 4; Page 148-149; 156pp; English.
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX partial sequence of pMal- $\alpha$ 3 (APETALA3) and ZFP $\alpha$ 3 DNA
XX SQ Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;
XX Query Match 97.8%; Score 3226.4; DB 4; Length 3300;
XX Best Local Similarity 98.6%; Pred. No. 0;
XX Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGTCGCAAAACCTTTTCGGGTATGCGATGATAGCCCGGAGAGA 60
```

Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCGGCGGTATGGCATGTAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTACAGGGTGGTGAATGTGAACACAGTAAACGTTATPACGATGTCGCAGAGTATGCG 120
Db 61 GTCAATTACAGGGTGGTGAATGTGAACACAGTAAACGTTATPACGATGTCGCAGAGTATGCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCGACGCAACGTTTCTGCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCGACGCAACGTTTCTGCGAAAA 180
Qy 181 CCGCGGAAAAAGTGAAGCGGAGTGGCGAGCTGAATTTACATTTCCACACCGCTGGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGGAGTGGCGAGCTGAATTTACATTTCCACACCGCTGGGCAC 240
Qy 241 AACAACTGGCGGCAAAACAGTGGTGTGATTTGGCGTTGCGACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTGGTGTGATTTGGCGTTGCGACCTCCAGTCTGGGCCCTGC 300
Qy 301 ACGCGCGTCGCAAAATTTGTCGGCGGATTTAAATCTCGCGCGGATCAACTGGGTGGCAGCG 360
Db 301 ACGCGCGTCGCAAAATTTGTCGGCGGATTTAAATCTCGCGCGGATCAACTGGGTGGCAGCG 360
Qy 361 TGGTGGTGTGATGGTGAACACGAGCGGCTGAAGCCCTGTAAAGCGCGGTGCAAAATC 420
Db 361 TGGTGGTGTGATGGTGAACACGAGCGGCTGAAGCCCTGTAAAGCGCGGTGCAAAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTCCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTCCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTCCGATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAAGTTCTGTCTCGG 660
Db 601 TGGTCCGATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAAGTTCTGTCTCGG 660
Qy 661 CCGCTGTGCTGTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CCGCTGTGCTGTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Qy 721 CCGAACGGGAAGCGACTGAGTGCCATGTCGGTTCGATTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAACGGGAAGCGACTGAGTGCCATGTCGGTTCGATTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Qy 841 TCGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATAG 900
Db 841 TCGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACCATCAAAAGGATTTTC 960
Qy 961 GCGTGTGGGCGAAACACGCGTGAGCGCTGTGTCGAATCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCGTGTGGGCGAAACACGCGTGAGCGCTGTGTCGAATCTCTCAGGCGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCTCATCTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCATCTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CCGAAACCGCCTCTCCCGCGGCTTGGCCGATTCATTAATGAGCTGGGACGACAGGTTT 1140

Db 1081 CCGAAACCGCCTCTCCCGCGGCTTGGCCGATTCATTAATGAGCTGGCAGCAGAGGTTT 1140
Qy 1141 CCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGTAATCACTGCGATTAATTCG 1320
Db 1261 TCAGGACCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGTAATCACTGCGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA 1500
Qy 1501 GCACCTTCCAAACAGGACCATAGATTATGAAAACTGAAGAGGTAATCTGGTAATCTGG 1560
Db 1501 GCACCTTCCAAACAGGACCATAGATTATGAAAACTGAAGAGGTAATCTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAATTCACAGGTT 1680
Qy 1681 CCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Db 1681 CCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGGAACAAAGGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGGAACAAAGGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGGTTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGGTTTATCGGTTCAAG 2040
Qy 2041 TATGAAAAACGGCAAGTACGACATTAAGAACGCTGGCGCTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGAACGCTGGCGCTGGATAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGAACCTTCTCGGTTGAACCTGATTTAAAAACAAACACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGAACCTTCTCGGTTGAACCTGATTTAAAAACAAACACATGAATGCGAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220

```
QY 2221 GCATGGTCCACATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCGAACATCCAAACCGTTCGTTGGCGTGTGAGCGAGGATTAACGCGCCAGT 2340
Db 2281 AAGGTCGAACATCCAAACCGTTCGTTGGCGTGTGAGCGAGGATTAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
QY 2461 TTGGGAAAGATCCACGATTTGCGGCCACCATGGAAGACGCCCGAAGAGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGATTTGCGGCCACCATGGAAGACGCCCGAAGAGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCGGTGCGTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCGGTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
QY 2641 AACACACACACATACAAATAACAAACCTCGGATCGAGGAGGAAGATTCAGATTC 2700
Db 2641 AACACACACACATACAAATAACAAACCTCGGATCGAGGAGGAAGATTCAGATTC 2700
QY 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTTCTCAGAGCTCTCAGAGCTCTCAGCTGTGCGCCACGAGTACC 2820
Db 2761 CCGGAATGTGTAAGTCTTCTCAGAGCTCTCAGAGCTCTCAGCTGTGCGCCACGAGTACC 2820
QY 2821 ACGGTTGAAAACCGTATAATGCCAGAGTCCGCAAAATCTTTAGCAGTCCAGAAC 2880
Db 2821 ACGGTTGAAAACCGTATAATGCCAGAGTCCGCAAAATCTTTAGCAGTCCAGAAC 2880
QY 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGTCCAGAAATGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGTCCAGAAATGGC 2940
QY 2941 AAGTCTTCTCTCGGTGTGAACAATCTCGTCCGGCAACCAAGTACTCACCGGGGAGAAG 3000
Db 2941 AAGTCTTCTCTCGGTGTGAACAATCTCGTCCGGCAACCAAGTACTCACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTCTCGGATGGTAAGTCTTCTAGCGCGCAGGATACCTGGTCCGCCAC 3060
Db 3001 CCCTATGCTTCTCGGATGGTAAGTCTTCTAGCGCGCAGGATACCTGGTCTAGACAC 3060
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCGCCAGAGTCCGCAAAATCTTTAGC 3120
Db 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCGCCAGAGTCCGCAAAATCTTTAGC 3120
QY 3121 CAGCGCGCCACCTGGCGAGCCATCAAGCACTCATATCTGGCGAGAGGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTGGCGAGCCATCAAGCACTCATATCTGGCGAGAGGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCGGTCTGACAAATCTCGTCCGGCACCAAGCTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCGGTCTGACAAATCTCGTCCGGCACCAAGCTACTCAC 3240
QY 3241 ACCGGTAAAAAATACTAGTGGCCAGCGGCCAGTACCCGTACGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATACTAGTGGCCAGCGGCCAGTACCCGTACGAGTTCGGGACTACGCT 3300
```

```
RESULT 4
RAD11591
ID AAD11591 standard; DNA; 3300 BP.
XX
AC AAD11591;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m4 and ZFPm4 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2719..3270
FT /tag= a
FT /product= "ZFPm4 protein"
FT /transl_except= (pos:3046..3048, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /tag= m
FT /bound_moiety= "F6-b2 primer"
XX
PN WO200152620-A2.
XX
PD 26-JUL-2001.
XX
PP 19-JAN-2001; 2001WO-US001817.
XX
PR 21-JAN-2000; 2000US-0177468P.
XX
PR 21-JUL-2000; 2000US-0062089P.
XX
PA (SCRI ) SCRIPPS RES INST.
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX Barbas CF, Stege JT, Guan X, Dalmia B;
PI
```


XX WPI; 2001-465325/50.
DR P-PSDB; AAE06004.
XX
PT New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 145-147; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m4 and ZFPm4 DNA
XX
SQ Sequence 3300 BP; 849 A; 865 C; 880 G; 706 T; 0 U; 0 Other;
Query Match 97.4%; Score 3215.2; DB 4; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 1 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGCGATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGCGATGATAGCGCCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGGTGAATGTGAACAGTAACCTTATACGATCTCCGAGAGTATCGG 120
DB 61 GTCATTCAGGGTGGTGAATGTGAACAGTAACCTTATACGATCTCCGAGAGTATCGG 120
QY 121 GTGCTCTTATCAGACCGGTTTCCCGGTGGTGAACAGCGCCAGCCACGCTTTCTGCGAAAA 180
DB 121 GTGCTCTTATCAGACCGGTTTCCCGGTGGTGAACAGCGCCAGCCACGCTTTCTGCGAAAA 180
QY 181 CCGGGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
DB 181 CCGGGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
QY 241 AACRACTGCGGGCAACAGTCTGTTGCTGATTTGGCGTTTGCACCTCCAGTCTGGCCCTGC 300
DB 241 AACRACTGCGGGCAACAGTCTGTTGCTGATTTGGCGTTTGCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGCGGGCGATTAATACTCTCGCGCGGATCAACTGGGTGGCAGCG 360
DB 301 ACGCGCGCTCGCAAAATTTGCGGGCGATTAATACTCTCGCGCGGATCAACTGGGTGGCAGCG 360
QY 361 TGGTGGTTCGATGTAGAACGAAGCGCGTCCGAGCTGTAAAGCGCGGTGCAACAATC 420
DB 361 TGGTGGTTCGATGTAGAACGAAGCGCGTCCGAGCTGTAAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
DB 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTGCATTTGGGTACACGAAATATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTGCATTTGGGTACACGAAATATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG

DB 601 TGGTGCATTTGGGTACACGAAATATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CCGCTCTCGCTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CCGCTCTCGCTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCAATGCAAAATCTCTGA 780
DB 721 CCGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCAATGCAAAATCTCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTTGCACACGATCAGATGGCGCTGGGCGCAA 840
DB 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTTGCACACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGCCCATTAACGAGTCCCGGCTGGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGCCCATTAACGAGTCCCGGCTGGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGTACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAACAGGATTTTC 960
DB 901 ACGTACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGCAGGTTT 1140
DB 1081 CGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGCAGGTTT 1140
QY 1141 CCGCACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 1141 CCGCACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGTGCAACCAATGTTCTGGCG 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGTGCAACCAATGTTCTGGCG 1260
QY 1261 TCAGGACCGCATCGAAGCTGTGTTATGCTGTGCGAGCTCGTAAATCATCTGATTAATTCG 1320
DB 1261 TCAGGACCGCATCGAAGCTGTGTTATGCTGTGCGAGCTCGTAAATCATCTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCGACTCCCGTCTCGGATTAATGTTTTTGGCGCGACATCAATAACGGTT 1380
DB 1321 TGTGCTCAAGCGGCGACTCCCGTCTCGGATTAATGTTTTTGGCGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTAGTGTTCACGA 1500
DB 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTAGTGTTCACGA 1500
QY 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAATACTGGTAATCTGG 1560
DB 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAATACTGGTAATCTGG 1560
QY 1561 ATTAACCGGCGATTAAGGCTATAAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
DB 1561 ATTAACCGGCGATTAAGGCTATAAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCACAGGTT 1680
QY 1681 GCGGCAATCTGGGATGGCGCTGACATTAATCTCTGGGCGACAGCGCTTCTGGCTGCTAC 1740
DB 1681 GCGGCAATCTGGGATGGCGCTGACATTAATCTCTGGGCGACAGCGCTTCTGGCTGCTAC 1740

```
QY 1741 GCTCAATCTGGCTCTGTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTCTGTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTAGCTTTAACCGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTTAACCGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACGCTATTACCTGGCGCTGATGCTGCTGAGCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGCTATTACCTGGCGCTGATGCTGCTGAGCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGATGAGCGGTGGAATAAGCTGGCGGCAAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGATGAGCGGTGGAATAAGCTGGCGGCAAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACACATGAATGCAACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACACATGAATGCAACACCGATTAC 2160
QY 2161 TCCATCGGAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACATCAACCGGCCGTGG 2220
Db 2161 TCCATCGGAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACATCAACCGGCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGTTAAACGGTACTGCCACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGTTAAACGGTACTGCCACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTGAAAACCTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTGAAAACCTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATGGAACACGCCCAAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATGGAACACGCCCAAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCAAGATGTCGGTCTTCTGATGAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2580
Db 2521 CCGAACATCCGCAAGATGTCGGTCTTCTGATGAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2580
QY 2581 GCCAGCGGTGTCAGACTGTGAGTAAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGTCAGACTGTGAGTAAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTCGCGCCAGGCGCCCTGAGCCCGGGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCTCTTCTCTGTCGCGCCAGGCGCCCTGAGCCCGGGGAGAGCCCTATGCTGT 2760
QY 2761 CCGGAATGTGGTAAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCACCAAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCACCAAGCGTACCCAC 2820
```

```
QY 2821 ACGGGTGAAAACCGTATAAATGCCAGAGTGGCGGAAAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGGTGAAAACCGTATAAATGCCAGAGTGGCGGAAAATCTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCGCCATCAACCGCACTCATCTGCGGAGAAAGCCATACAAATGTTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCGCCATCAACCGCACTCATCTGCGGAGAAAGCCATACAAATGTTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTCACACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTCTCGGAAATGTGTAAGTCTTCTGAGCCGAGCGGTAACCTGGTGGCGCAC 3060
Db 3001 CCCTATGCTTCTCGGAAATGTGTAAGTCTTCTGAGCCGAGCGGTAACCTGGTGGCGCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAACCGTATAAATGCCAGAGTGGCGGAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAACCGTATAAATGCCAGAGTGGCGGAAAATCTTTTAGC 3120
QY 3121 CAGCGCGCCACCTGGCCAGCCATCAACGCACTCATCTGCGGAGAAAGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTGGCCAGCCATCAACGCACTCATCTGCGGAGAAAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGCGAAGTCTTTCTCTCGGTCTGCAATCTCTGCGGACCAACGTAATCTCAC 3240
Db 3181 CCAGAAATGTGCGAAGTCTTTCTCTCGGTCTGCAATCTCTGCGGACCAACGTAATCTCAC 3240
QY 3241 ACCGGTAAAAAAACTAGTGGCCAGGCGCGCCAGTACCCGTAACGCTTCGGGACTAGCT 3300
Db 3241 ACCGGTAAAAAAACTAGTGGCCAGGCGCGCCAGTACCCGTAACGCTTCGGGACTAGCT 3300

RESULT 5
AAD11590
ID AAD11590 standard; DNA; 3300 BP.
XX
AC AAD11590;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m3 and ZFPm3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
Key Location/Qualifiers
CDS 2719..3270
FT /tag= a
FT /product= "ZFPm3 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
```

```
FT      /*tag= h
FT      /bound_moiety= "F4-f2 primer"
FT      3022..3102
FT      /*tag= i
FT      /bound_moiety= "F4-f1 primer"
FT      3076..3141
FT      /*tag= j
FT      /bound_moiety= "F5-b primer"
FT      3119..3192
FT      /*tag= k
FT      /bound_moiety= "F5-f primer"
FT      3168..3225
FT      /*tag= l
FT      /bound_moiety= "F6-b1 primer"
FT      3205..3273
FT      /*tag= m
FT      /bound_moiety= "F6-b2 primer"
XX
PN      WO200152620-A2.
XX
XX
PD      26-JUL-2001.
XX
XX      19-JAN-2001; 2001WO-US001817.
XX
XX      21-JAN-2000; 2000US-0177468P.
XX      21-JUL-2000; 2000US-00620897.
XX
PA      (SCRI ) SCRIPPS RES INST.
PA      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX      Barbac CF, Stege JT, Guan X, Dalmia B;
XX
XX      WPI; 2001-465325/50.
XX      P-PSDB; AA06003.
XX
PT      New zinc finger proteins, useful for modulating or regulating gene
PT      expression and metabolic pathways in plants, e.g. for treating in the
PT      plant cells a disorder that is associated with abnormal expression of the
PT      target gene.
XX
PS      Example 4; Page 143-145; 156pp; English.
XX
XX      The patent discloses methods and compositions to modulate the expression
XX      of a target gene in plant cells. The method involves providing plant
XX      cells with a zinc finger protein (ZFP) which is capable of specifically
XX      binding to a target nucleotide sequence or its complementary strand
XX      within a target gene and allowing the ZFP binding to the target
XX      nucleotide sequence, where the expression of the target gene in the plant
XX      cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX      for modulating or regulating gene expression and metabolic pathways in
XX      plants. The ZFP, fusion proteins and methods are useful in plant and
XX      agricultural technology. The method is useful particularly for treating a
XX      disorder in the plant cells, where the disorder is associated with
XX      abnormal expression of the target gene. The present DNA sequence is the
XX      Partial sequence of pMal-m3 and ZFPm3 DNA
SQ      Sequence 3300 BP; 845 A; 878 C; 877 G; 700 T; 0 U; 0 Other;
Query Match          97.2%; Score 3207.2; DB 4; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY      1   CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA 60
DB      |
DB      1   CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA 60
QY      61  GTCAATTGAGGTGTGAATGTGAACCCAGTAAACGTTATACGATTCGCAGAGTATGCCG 120
DB      |
DB      61  GTCAATTGAGGTGTGAATGTGAACCCAGTAAACGTTATACGATTCGCAGAGTATGCCG 120
QY      121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCCAGGCCAGCCAGCTTCTCGGAAA 180
DB      |
DB      121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCCAGGCCAGCCAGCTTCTCGGAAA 180
```

```
QY      181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
DB      |
DB      181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
QY      241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB      |
DB      241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY      301 ACGGCCGTCGCAAAATTTGTTCGCGCGGATTAATCTCGCGCGATCAACTCGGTGCGCAGCG 360
DB      |
DB      301 ACGGCCGTCGCAAAATTTGTTCGCGCGGATTAATCTCGCGCGATCAACTCGGTGCGCAGCG 360
QY      361 TGGTGGTGTGATGGTAGAACGAGCGCGTCTGAAGCCTGTAAGCGCGCGGTGCACAATC 420
DB      |
DB      361 TGGTGGTGTGATGGTAGAACGAGCGCGTCTGAAGCCTGTAAGCGCGCGGTGCACAATC 420
QY      421 TTCTCGCGCAACGCGTCACTGGGCTGATCACTAACTATCCGCTGGATCACCAGGATGCCA 480
DB      |
DB      421 TTCTCGCGCAACGCGTCACTGGGCTGATCACTAACTATCCGCTGGATCACCAGGATGCCA 480
QY      481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
DB      |
DB      481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
QY      541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGCGTGGAGCATC 600
DB      |
DB      541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGCGTGGAGCATC 600
QY      601 TGGTCGCAATGGGTCCAGCAAAATCGCGCTGTTAGCGGGGCCATTAGTTCTGCTCGG 660
DB      |
DB      601 TGGTCGCAATGGGTCCAGCAAAATCGCGCTGTTAGCGGGGCCATTAGTTCTGCTCGG 660
QY      661 CGCGTCTCGCTCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB      |
DB      661 CGCGTCTCGCTCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY      721 CGGAAACGGGAAGGCGATGGAGTGCATGTCGGTTCGTTTCAACAAACCAATGCAAAATGCTGA 780
DB      |
DB      721 CGGAAACGGGAAGGCGATGGAGTGCATGTCGGTTCGTTTCAACAAACCAATGCAAAATGCTGA 780
QY      781 ATGAGGGCATGTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
DB      |
DB      781 ATGAGGGCATGTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
QY      841 TGGCGCCATTACCGAGTCCCGGCTGCGCTGGTGGCGATATCTCGGTAGTGGGATAGC 900
DB      |
DB      841 TGGCGCCATTACCGAGTCCCGGCTGCGCTGGTGGCGATATCTCGGTAGTGGGATAGC 900
QY      901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATGCAAAAGGATTTTC 960
DB      |
DB      901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATGCAAAAGGATTTTC 960
QY      961 GCCTGTGGGGCAACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGCCACGCGGTGA 1020
DB      |
DB      961 GCCTGTGGGGCAACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGCCACGCGGTGA 1020
QY      1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
DB      |
DB      1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
QY      1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCATTAAATGCAAGTGGCACGACAGGTTT 1140
DB      |
DB      1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCATTAAATGCAAGTGGCACGACAGGTTT 1140
QY      1141 CCCGATCGAAAGCGGCGAGTGAGGCGAACCGCAATTAATGAGTGTAGTCTCACTCATTTAG 1200
DB      |
DB      1141 CCCGATCGAAAGCGGCGAGTGAGGCGAACCGCAATTAATGAGTGTAGTCTCACTCATTTAG 1200
QY      1201 GCACAATTCATGTTTCACAGCTTATCATGACGTGACGCGGTGACCAATGCTTCTGGCG 1260
DB      |
DB      1201 GCACAATTCATGTTTCACAGCTTATCATGACGTGACGCGGTGACCAATGCTTCTGGCG 1260
```

QY	1261	TCAGGAGCCATCGAAGCTGTGGTATGCTGTGAGGTGTAATCACTGCATATTCG	1320
Db	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTGTAATCACTGCATATTCG	1320
QY	1321	TGTCGCTCAAGCGGCACCTCCGCTTCGGATATGTTTTTGGCGCGACATCAATACGGTT	1380
Db	1321	TGTCGCTCAAGCGGCACCTCCGCTTCGGATATGTTTTTGGCGCGACATCAATACGGTT	1380
QY	1381	CTGGCAATATTTCTGAATGAGCTGTGCACATTAATCATCGGCTCGTATATGTTGGA	1440
Db	1381	CTGGCAATATTTCTGAATGAGCTGTGCACATTAATCATCGGCTCGTATATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTAGGTGTTTTACGA	1500
Db	1441	ATTGTGAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTAGGTGTTTTACGA	1500
QY	1501	GCACCTTACCACAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAACTCTGG	1560
Db	1501	GCACCTTACCACAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAACTCTGG	1560
QY	1561	ATTAAACGCGATAAAGGCTATAACGCTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGCGATAAAGGCTATAACGCTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGTT	1680
Db	1621	ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGTT	1680
QY	1681	CGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTGGTGGCTAC	1740
Db	1681	CGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCGCTGTGGCTGAAATCAACCCGACAAAGCGTTCCAGGACAACTGTAT	1800
Db	1741	GCTCAATCTGCGCTGTGGCTGAAATCAACCCGACAAAGCGTTCCAGGACAACTGTAT	1800
QY	1801	CCGTTTACCTGGATGCGGTACGTTTACAAACGGAAGCTGATGCTTACCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGATGCGGTACGTTTACAAACGGAAGCTGATGCTTACCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTATTAACAAGATCTGTCGCGAACCGCCAAACCTCGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTATTAACAAGATCTGTCGCGAACCGCCAAACCTCGGAA	1920
QY	1921	GAGATCCGCGCGTGGATTAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGCGCGTGGATTAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTCGCGCTGATGCTGCTGACGGGGTTATGCTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTCGCGCGCTGATGCTGCTGACGGGGTTATGCTTCAAG	2040
QY	2041	TATGAAACCGCAAGTACGACATTAAGAAGTGGCGGTGGATACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAACCGCAAGTACGACATTAAGAAGTGGCGGTGGATACGCTGGCGGAAAGCG	2100
QY	2101	GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGAGACACCGATTAC	2160
QY	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCGGTGG	2220
Db	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCGGTGG	2220
QY	2221	GCATGCTCAACATCGACACCGAAGTGAATTTATGTTAAGCGTACTGCCGACCTTC	2280
Db	2221	GCATGCTCAACATCGACACCGAAGTGAATTTATGTTAAGCGTACTGCCGACCTTC	2280
QY	2281	AAGGTCACACATCCAAACCGTTGCTGGCGTGTAGCGCAGGTATTAACGCGCGCAGT	2340
Db	2281	AAGGTCACACATCCAAACCGTTGCTGGCGTGTAGCGCAGGTATTAACGCGCGCAGT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGATTCTCTCGAAAACTATCTGCTGACTGATGAAGGCTG	2400

Db	2341	CCGAACAAAGAGCTGGCAAAAGATTCTCGAAAACTATCTGCTGACTGATGAAGGCTG	2400
QY	2401	GAAGCGTTTAATAAAGACAAACCGCTGGTGGCGTAGCTGAACTTTACGAGGAAGAG	2460
Db	2401	GAAGCGTTTAATAAAGACAAACCGCTGGTGGCGTAGCTGAACTTTACGAGGAAGAG	2460
QY	2461	TTGGCGAAAGATCCACGTTATTCGCGCACCATGGAACCGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTTATTCGCGCACCATGGAACCGCCAGAAAGGTGAATCATG	2520
QY	2521	CCGAACATCCGCGAGATGTCGCTTTCTGTTATGCGGTGCTGCTGGGTGATCAACGCC	2580
Db	2521	CCGAACATCCGCGAGATGTCGCTTTCTGTTATGCGGTGCTGCTGGGTGATCAACGCC	2580
QY	2581	GCACGCGTCTGACAGCTGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCACGCGTCTGACAGCTGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
QY	2641	AACAACAAACAACTAACAATTAACAACCTCGGGATCGAGGAAGGATTCAGAAATTC	2700
Db	2641	AACAACAAACAACTAACAATTAACAACCTCGGGATCGAGGAAGGATTCAGAAATTC	2700
QY	2701	GGATCTCTTCTCTGTGGCCAGGCGGCCCTCGAGCCCGGGGAGAAAGCCCTATGCTGT	2760
Db	2701	GGATCTCTTCTCTGTGGCCAGGCGGCCCTCGAGCCCGGGGAGAAAGCCCTATGCTGT	2760
QY	2761	CCGGAATGTTGTAAGTCTTCTCTGAGAGCTCTCACTGTGTGCGCACACAGGTACCCAC	2820
Db	2761	CCGGAATGTTGTAAGTCTTCTCTGAGAGCTCTCACTGTGTGCGCACACAGGTACCCAC	2820
QY	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCCAGTCCAGCAAC	2880
QY	2881	CTGCTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC	2940
Db	2881	CTGCTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC	2940
QY	2941	AAGTCTTCTCTCGTCTGACAACTCTGTCGCGCACCAAGCTACTCACACGGGGAGAAG	3000
Db	2941	AAGTCTTCTCTCGTCTGACAACTCTGTCGCGCACCAAGCTACTCACACGGGGAGAAG	3000
QY	3001	CCCTATGCTTGTCCGGAATGTGGTAAAGTCTCTCAGCGCGCAGCGATAACCTGTGCGCCAC	3060
Db	3001	CCCTATGCTTGTCCGGAATGTGGTAAAGTCTCTCAGCGCGCAGCGATAACCTGTGCGCCAC	3060
QY	3061	CAGCGTACCCACACCGGGTGAATAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTTAGT	3120
Db	3061	CAGCGTACCCACACCGGGTGAATAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTTAGT	3120
QY	3121	CAGGCGCGCCACCTGGCCAGCCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGT	3180
Db	3121	CAGGCGCGCCACCTGGCCAGCCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGT	3180
QY	3181	CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGCAAACTCTGTCGCGCACCAAGCTACTCAC	3240
Db	3181	CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGCAAACTCTGTCGCGCACCAAGCTACTCAC	3240
QY	3241	ACCGGTAAAAAACTAGTGGCCAGGCGCGGCGAGTACCCGTACGACGTTCCGGACTACGCT	3300
Db	3241	ACCGGTAAAAAACTAGTGGCCAGGCGCGGCGAGTACCCGTACGACGTTCCGGACTACGCT	3300

RESULT 6

AAH25681

ID AAH25681 standard; DNR; 8101 BP.

XX

AC AAH25681;

XX

DT 05-SEP-2001 (first entry)

XX

DE Nucleotide sequence of human Cdc25C phosphatase DNA.

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
KW ss.
XX
XX Homo sapiens.
XX WO200144467-A2.
XX 21-JUN-2001.
XX 13-DEC-2000; 2000WO-PR003496.
XX 14-DEC-1999; 99FR-00015722.
XX 30-MAY-2000; 2000FR-00006883.
XX 21-SEP-2000; 2000FR-00012008.
XX (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX Goubin-Gramatica F, Ducommun B, Prevost G;
XX WPI; 2001-398152/42.
XX New fusion protein of maltose-binding protein and Cdc phosphatase, useful
PT for identifying phosphatase modulators for regulating the cell cycle.
XX
XX Claim 7; Page 9-13; 56pp; French.
XX The present sequence encodes human Cdc25C phosphatase, and was expressed
CC in Escherichia coli. It is used to make fusion proteins with Escherichia
CC coli maltose binding protein (MBP). The specification describes fusion
CC proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
CC are involved in activation of cyclin-dependent kinases (CDK) that control
CC cell division. The fusion proteins are used to identify modulators of the
CC specified human Cdc phosphatases, potentially useful for regulation of
CC cell division. They are also for studying physiological or
CC physiopathological activities of Cdc phosphatase
XX
XX Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCGGAAGAGA 60
DB 1 CGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCGGAAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACCGTTATACGATGTCGAGAGTATGCCG 120
DB 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACCGTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCTCTATCAGACCGTTTCCGGGTGGTGAACACAGGCGCCAGCCAGTTCTGCGAATA 180
DB 121 GTGTCTCTATCAGACCGTTTCCGGGTGGTGAACACAGGCGCCAGCCAGTTCTGCGAATA 180
QY 181 CGCGGGAAAAAGTGAAGCGCGGATGGCGAGTCGAATTAACATTCGCAACCGGTGGCAC 240
DB 181 CGCGGGAAAAAGTGAAGCGCGGATGGCGAGTCGAATTAACATTCGCAACCGGTGGCAC 240
QY 241 AACAACTGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 AGCGCGGTGCGAAATGTGCGGGGATTAATCTCGCGCCGATCAACTGGGTGGCCAGCG 360
DB 301 AGCGCGGTGCGAAATGTGCGGGGATTAATCTCGCGCCGATCAACTGGGTGGCCAGCG 360
QY 361 TGGTGTGTGTCGATGTGTAAGACGCGGTGTAAGCCCTGTAAGCGCGGTGCAAAATC 420
DB 361 TGGTGTGTGTCGATGTGTAAGACGCGGTGTAAGCCCTGTAAGCGCGGTGCAAAATC 420
QY 421 TTCTCGCGCAACCGGTGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACAGGATGCCA 480

Db 421 TTCTCGCGCAACCGGTGTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGATGCCA 480
QY 481 TTGTGTGTGGAAGCTGCTGACCTAAATGTTCCGGCGTTATTTCTTGTATCTCTGTACCCAGA 540
Db 481 TTGTGTGTGGAAGCTGCTGACCTAAATGTTCCGGCGTTATTTCTTGTATCTCTGTACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCAATTCGGTCCAGCAAAATTCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCAATTCGGTCCAGCAAAATTCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
QY 841 TGGCGCCATTAACGAGTCCGGCTGCGGTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGCCATTAACGAGTCCGGCTGCGGTGGTGGCGATATCTCGGTAGTGGGATAG 900
QY 901 AGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCAATCAAAAGGATTTTC 960
Db 901 AGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCAATCAAAAGGATTTTC 960
QY 961 GCCTCTGCGGCAACACGAGTGGACCGCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTCTGCGGCAACACGAGTGGACCGCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAACCAACCTTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAACCAACCTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGAGCTGGCAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGAGCTGGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGTCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGTCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATTCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGATTAATTCG 1320
Db 1261 TCAGCAGCCATTCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGATTAATTCG 1320
QY 1321 TGTCCCTCAAGCGGCACTCCCGTTCGATATGTTTGGCGCGGACATCAATACGGTT 1380
Db 1321 TGTCCCTCAAGCGGCACTCCCGTTCGATATGTTTGGCGCGGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTATGAAAACTGAAGAAAGGTAATACTGGTAACTCTGG 1560

Db 1501 GCATCTCCAAACAGGACCAATAGATTATGAAAAATCGAAGAGGTAAACTGTTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAAGAGAAATTCCTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAAGAGAAATTCCTCCACAGGTT 1680
Qy 1681 GCGCAACTGCGGATGGCCCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGGCCCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCGCCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCGCCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGTATACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGTATACCGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGTCGCAACCGCCCAAAACCTGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGTCGCAACCGCCCAAAACCTGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTATCTTACCTGCGCGCTGATTGCTGCTGACGGGGTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTATCTTACCTGCGCGCTGATTGCTGCTGACGGGGTATGCTTCAAG 2040
Qy 2041 TATGAAACGCGCAAGTACGACATTAAAGACGTGGCGCTGGATACGCTGGCGCAAGCG 2100
Db 2041 TATGAAACGCGCAAGTACGACATTAAAGACGTGGCGCTGGATACGCTGGCGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAACAAACACACATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAACAAACACACATGAATGACACACCGATTAC 2160
Qy 2161 TCCATCGAAGAGTGCCTTTAATAAGGCGAAACAGCGATGACATCAACGGCCCGTGG 2220
Db 2161 TCCATCGAAGAGTGCCTTTAATAAGGCGAAACAGCGATGACATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCACCTTC 2280
Qy 2281 AAGGGTCAACCATCAACCGCTGCGTGGCGTGTGAGCGCAGGTATTACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAACCGCTGCGTGGCGTGTGAGCGCAGGTATTACGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCGTGTGAGCGCAGGTATTACGCCGCCAGT 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCGTGTGAGCGCAGGTATTACGCCGCCAGT 2460
Qy 2461 TTGCGGAAAGATCCACGTATTGCGCCACCATGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTATTGCGCCACCATGAAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGGCTTTCTGTTATCCGTCGGTATCGGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGGCTTTCTGTTATCCGTCGGTATCGGGTGATCAAGCC 2580
Qy 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACGGCGCAGACTAATTCAGGTCG 2640
Db 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACGGCGCAGACTAATTCAGGTCG 2640

Qy 2641 AACAAACAACATATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAGAATTC 2700
Db 2641 AACAAACAACATATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAGAATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709
RESULT 7
ABL53238
ID ABL53238 standard; DNA; 8101 BP.
XX ABL53238;
AC ABL53238;
XX 17-JUN-2002 (first entry)
DT Nucleotide sequence of pMAL-Hs Cdc25C.
XX Cdc25C phosphatase; amine; Cdc25 phosphatase; spontaneous alopecia;
XX alopecia; proliferative disease; parasitic disease; viral infection;
XX neurodegenerative disease; myopathy; angiogenesis; psoriasis; restenosis;
XX vitamin K; ss.
OS Synthetic.
OS Homo sapiens.
XX Key Location/Qualifiers
FT CDS 1528..2677
FT /*tag= b
FT /note= "encodes maltose binding protein (MBP)"
FT CDS 2713..4134
FT /*tag= a
FT /note= "ORF of human Cdc25C"
XX FR2812198-A1.
XX 01-FEB-2002.
XX 28-JUL-2000; 2000FR-00009900.
XX 28-JUL-2000; 2000FR-00009900.
XX (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX Prevost G, Brezak PMC, Galcera CMO, Thuriereau C, Goubin GF;
XX Ducommun B;
XX WPI; 2002-166066/22.
XX cdc25 Phosphatase inhibiting medicaments, useful e.g. for treating tumor
XX diseases, viral infections, neurodegenerative disease or alopecia,
XX containing new or known aryl-substituted amines.
XX Disclosure; Page 18-22; 39pp; French.
XX The present sequence represents pMAL-HsCdc25C, a plasmid encoding human
XX Cdc25C phosphatase. The specification describes the use of aryl-
XX substituted secondary or tertiary amines for the production of
XX medicaments for inhibiting Cdc25 phosphatases, especially Cdc25C
XX phosphatases. The amines of the invention are used for treating
XX spontaneous alopecia or alopecia induced by exogenous products or
XX radiation. They are also used for treating tumoral or non-tumoral
XX proliferative diseases, parasitic diseases, viral infections,
XX neurodegenerative diseases or myopathy. Examples of non-tumoral
XX proliferative diseases are e.g. angiogenesis, psoriasis or restenosis.
XX The amines are also useful for inhibiting the proliferation of
XX microorganisms (especially yeasts) and in the treatment of all disorders
XX conventionally treated with vitamin K or its derivatives
XX Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match 82.0%; Score 2705.8; DB 6; Length 8101; Best Local Similarity 99.9%; Pred. No. 0; Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
Qy	1	CCGACACCATCGAATGGTGGCAAAACCTTTTCGGGTATGCGCATGATAGCGCCCGGAAGAGA	60						
Db	1	CCGACACCATCGAATGGTGGCAAAACCTTTTCGGGTATGCGCATGATAGCGCCCGGAAGAGA	60						
Qy	61	GTCAATTACAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCCAGAGTATGCCG	120						
Db	61	GTCAATTACAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCCAGAGTATGCCG	120						
Qy	121	GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTCCCAACCGGTGGCAC	180						
Db	121	GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTCCCAACCGGTGGCAC	180						
Qy	181	CGCGGAAATGTGAACGCGGATGGCGGAGCTGAATTAACATCCCAACCGGTGGCAC	240						
Db	181	CGCGGAAATGTGAACGCGGATGGCGGAGCTGAATTAACATCCCAACCGGTGGCAC	240						
Qy	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
Db	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
Qy	301	ACGGCCGTCCGAATTTGTGGCGGATTAATCTCGCGCGATCAACTGGGTGCGACG	360						
Db	301	ACGGCCGTCCGAATTTGTGGCGGATTAATCTCGCGCGATCAACTGGGTGCGACG	360						
Qy	361	TGTTGGTGTGATGTAAGCAAGCGCGTGAAGCCTGTAAAGCGCGGTGCAATC	420						
Db	361	TGTTGGTGTGATGTAAGCAAGCGCGTGAAGCCTGTAAAGCGCGGTGCAATC	420						
Qy	421	TTCTCGCCCAACCGCTCAGTGGGCTGATCAATTAATCTCCGTGATGACCAAGTGGCA	480						
Db	421	TTCTCGCCCAACCGCTCAGTGGGCTGATCAATTAATCTCCGTGATGACCAAGTGGCA	480						
Qy	481	TGCTGTGGAGTGCCTGCAATTAATGTTCGGCGTTATTTCTTGATGTCTGACGAGA	540						
Db	481	TGCTGTGGAGTGCCTGCAATTAATGTTCGGCGTTATTTCTTGATGTCTGACGACGA	540						
Qy	541	CACCATCAACAGTATTTCTCCATGAAGCGGTACGCGCTGGCGGTGGAGCATC	600						
Db	541	CACCATCAACAGTATTTCTCCATGAAGCGGTACGCGCTGGCGGTGGAGCATC	600						
Qy	601	TGTTGGCATTTGGTCAACAGAAATCGGCTGTTAGCGGGCCCAATAAGTCTGTCTGG	660						
Db	601	TGTTGGCATTTGGTCAACAGAAATCGGCTGTTAGCGGGCCCAATAAGTCTGTCTGG	660						
Qy	661	CGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Qy	721	CGGAACGGAGGCGACATGGAGTGCATGTCGGTTCACAAACCATGCAATGCTGA	780						
Db	721	CGGAACGGAGGCGACATGGAGTGCATGTCGGTTCACAAACCATGCAATGCTGA	780						
Qy	781	ATGAGGGCATCTTCCCACTGCGATGCTGGTGGCCAAACGATCAGATGGCGCTGGCGGCA	840						
Db	781	ATGAGGGCATCTTCCCACTGCGATGCTGGTGGCCAAACGATCAGATGGCGCTGGCGGCA	840						
Qy	841	TGCGGCCATTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATAG	900						
Db	841	TGCGGCCATTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATAG	900						
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACATCAACAGGATTTTC	960						
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACATCAACAGGATTTTC	960						
Qy	961	GCCTGCTGGGCAACACAGCGTGGACCGCTTCTCCTCACTCTCAGGCGCAGGCGGTGA	1020						
Db	961	GCCTGCTGGGCAACACAGCGTGGACCGCTTCTCCTCACTCTCAGGCGCAGGCGGTGA	1020						
Qy	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080						

Db	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080						
Qy	1081	CGCAAAACCGCTCTCTCCCGCGGTGGCCGATTCATTAATCAGCTGGCAGCAGCAGGTTT	1140						
Db	1081	CGCAAAACCGCTCTCTCCCGCGGTGGCCGATTCATTAATCAGCTGGCAGCAGCAGGTTT	1140						
Qy	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
Db	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGACCAATGCTTCGGCG	1260						
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGACCAATGCTTCGGCG	1260						
Qy	1261	TCAGGAGCCATCCGAAAGCTGTGATGGCTGTGAGCTGTCAGGTCGTAATCACTGCATAATTCG	1320						
Db	1261	TCAGGAGCCATCCGAAAGCTGTGATGGCTGTGAGCTGTCAGGTCGTAATCACTGCATAATTCG	1320						
Qy	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTTCGGCGCAGCATCAATAACGGTT	1380						
Db	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTTCGGCGCAGCATCAATAACGGTT	1380						
Qy	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGGA	1440						
Db	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGGA	1440						
Qy	1441	ATTGTGAGCGGATAACAATTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAGGA	1500						
Db	1441	ATTGTGAGCGGATAACAATTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAGGA	1500						
Qy	1501	GCATTTCAACCAAGGACCATAGATTATGAATACTGAAGAAAGTAAACTGGTAATCTGG	1560						
Db	1501	GCATTTCAACCAAGGACCATAGATTATGAATACTGAAGAAAGTAAACTGGTAATCTGG	1560						
Qy	1561	ATTAACCGCGATAAGGCTATAACGCTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620						
Db	1561	ATTAACCGCGATAAGGCTATAACGCTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620						
Qy	1621	ACCGAATTAAGTCAACCGTTGAGCATCCGATATACTGGAAGAGAAATTCCTCCAGGTT	1680						
Db	1621	ACCGAATTAAGTCAACCGTTGAGCATCCGATATACTGGAAGAGAAATTCCTCCAGGTT	1680						
Qy	1681	CGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Db	1681	CGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Qy	1741	GCTCAATCTGGCTGTTGGCTGAATCACCCCGGCAAGCGTTCCAGGACAGCTGTAT	1800						
Db	1741	GCTCAATCTGGCTGTTGGCTGAATCACCCCGGCAAGCGTTCCAGGACAGCTGTAT	1800						
Qy	1801	CCGTTTACCTGGGATGGCTACGTTTACAAACCGCAAGCTGATTGCTTACCCGATCGCTGT	1860						
Db	1801	CCGTTTACCTGGGATGGCTACGTTTACAAACCGCAAGCTGATTGCTTACCCGATCGCTGT	1860						
Qy	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980						
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980						
Qy	1981	CTGCAAGAAACGCTACTTCCACCTGGCGCTGATGCTGCTGAGGGGGTTATCGCTTCAAG	2040						
Db	1981	CTGCAAGAAACGCTACTTCCACCTGGCGCTGATGCTGCTGAGGGGGTTATCGCTTCAAG	2040						
Qy	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGGCGTGGATAAACGCTGGCGCGAAAGCG	2100						
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGGCGTGGATAAACGCTGGCGCGAAAGCG	2100						
Qy	2101	GGTCTGACCTTCTCGGTTGACCTGATTAATAAACAACATGAATGACAGACCCGATTAC	2160						
Db	2101	GGTCTGACCTTCTCGGTTGACCTGATTAATAAACAACATGAATGACAGACCCGATTAC	2160						


```
Db 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAATAAACAACACATGATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACACAGCAAGAGTGAATTATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACAGCAAGAGTGAATTATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCTGTCGGCTGCTGAGGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGTCGGCTGCTGAGGCGAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTGAGTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTGAGTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGTTGATGCGGTGCTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGTTGATGCGGTGCTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCAGGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCAGGCTCG 2640
QY 2641 AACACAAACAATAAACAATAACAACAACCTCGGATCGAGGGAAGATTTTCAAGATTC 2700
Db 2641 AACACAAACAATAAACAATAACAACAACCTCGGATCGAGGGAAGATTTTCAAGATTC 2700
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709
```

RESULT 8
AD023608

```
ID AD023608 standard; DNA; 7259 BP.
AC
XX
AC AD023608;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30del18(82-294aa) fusion protein.
XX
KW P30 antigen; Toxop30del13C; Toxop30del12C; Toxop30 MIX1;
KW MBP-Toxop30del13C(52-300aa); MBP-Toxop30del14C(52-294aa); MBP-Toxop30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
```

```
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
```

```
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
```

```
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
```

```
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; AD023609.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxop30del13C,
PT Toxop30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
XX Example 2; Fig 22; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxop30del13C, Toxop30del12C and Toxop30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxop30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxop30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxop30del13C(52-300aa), MBP-Toxop30del14C(52-294aa) and MBP-Toxop30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
XX Sequence 7259 BP; 1870 A; 1822 C; 1926 G; 1641 T; 0 U; 0 Other;
```

```
Query Match 81.5%; Score 2690.4; DB 12; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 CCACACCATCGAATGGTGCAGAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCACACCATCGAATGGTGCAGAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACCACTAGTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAACCACTAGTATACGATGTCGAGAGTATGCCG 120
QY 121 GTCTCTTTATCAGACCGTTTCCCGGTGTGAACCAAGCCAGCCACCGTTTTCGCGAAA 180
Db 121 GTCTCTTTATCAGACCGTTTCCCGGTGTGAACCAAGCCAGCCACCGTTTTCGCGAAA 180
QY 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCACCGCGTGGCAC 240
Db 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAACACAGTCTGCTGATTTGGCGTGTGCCACTCGAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAACACAGTCTGCTGATTTGGCGTGTGCCACTCGAGTCTGGCCCTGC 300
QY 301 ACGCGCGTGCAGAAATTTGTCGGCGCATTAATTTCTCGCGCCGATCAATCTGGGTGCGAGCG 360
Db 301 ACGCGCGTGCAGAAATTTGTCGGCGCATTAATTTCTCGCGCCGATCAATCTGGGTGCGAGCG 360
QY 361 TGGTGTGTGATGTTAGAACGAGCGGTGAAAGCCCTGTAAGCGCGGTGCACAATC 420
Db 361 TGGTGTGTGATGTTAGAACGAGCGGTGAAAGCCCTGTAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATTCATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATTCATCCGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATTTCCGGCGTATTTCTTGATGTCTGTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAATTTCCGGCGTATTTCTTGATGTCTGTGACACAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGAGCGGTACGCGACTGGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGAGCGGTACGCGACTGGGCGGTGGAGCATC 600
```

QY 601 TGGTCGCAATTGGGTCACACAGCAAAATCGCGCTGTAGCGGGGCCATTAAAGTTCTGTCTCGG 660
Db 601 TGGTCGCAATTGGGTCACACAGCAAAATCGCGCTGTAGCGGGGCCATTAAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTCTGGCTGGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTCTGGCTGGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGGGATGCTGGTTCACACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGGGATGCTGGTTCACACGATCAGATGGCGTGGCGCAA 840
QY 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTCTGTCGAATCTCTCAGGCGCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTCTGTCGAATCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGGCGCTTCCACTGGTGAAAGAAACACACCTGGCGGCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCGCTTCCACTGGTGAAAGAAACACACCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCAGCACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCAGCACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGAAACCGAATTAATGAGTGTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGAAACCGAATTAATGAGTGTAGTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTCAGACGTTATCATGACCTGCAGTGACCGTGCACCAATGCTTCGGG 1260
Db 1201 GCACAATTCATGTTTCAGACGTTATCATGACCTGCAGTGACCGTGCACCAATGCTTCGGG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGCGCACATCCCGTTCTGGATATGTTTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACATCCCGTTCTGGATATGTTTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTCGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTCGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCGTCCGTTTAGTGTTTTTCAAGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCGTCCGTTTAGTGTTTTTCAAGA 1500
QY 1501 GCACCTTCAACAAAGGACCATAGATATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTACGGGGATAAAGGCTATTAACCGTCTCGTGAAGTCGGTGAAGAAATTCAGAAAGAT 1620
Db 1561 ATTACGGGGATAAAGGCTATTAACCGTCTCGTGAAGTCGGTGAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACAGACCGCTTTGGTGGCTAC 1740

RESULT 9
ADO23613
ID ADO23613 standard; DNA; 7322 BP.
XX
AC ADO23613;

Db 1681 GCGCAAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACAA CGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTACAA CGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAA CAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAA CAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAAACAAAAACAATGATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAAACAAAAACAATGATGACAGACACCGATTAC 2160
QY 2161 TCCATTCGCAAGTGTGCTTTTAAAGCGAAACAGCGATGAACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATTCGCAAGTGTGCTTTTAAAGCGAAACAGCGATGAACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGCTCCACATCGACACAGCAAGTGAATTAATGCTGTAAACGCTGACCGACCTTC 2280
Db 2221 GCATGCTCCACATCGACACAGCAAGTGAATTAATGCTGTAAACGCTGACCGACCTTC 2280
QY 2281 AAGGTCACCAACCATCAAAACCGTTGTTGGCGTGTGAGCGAGGTATTTAAAGCGCCCACT 2340
Db 2281 AAGGTCACCAACCATCAAAACCGTTGTTGGCGTGTGAGCGAGGTATTTAAAGCGCCCACT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTTAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCAAGAGATCCACGTTATTTGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCAAGAGATCCACGTTATTTGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTGTAAGCCCTTGAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTGTAAGCCCTTGAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAATAAACAATAAACAACCACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACCAACAAACAATAAACAATAAACAACCACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

XX 01-JUL-2004 (first entry)
 XX DNA encoding MBP-ToxoP30del110(52-284aa) fusion protein.
 XX
 XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
 KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; da; gene.
 XX
 XX Toxoplasma gondii.
 OS Synthetic.
 OS
 XX US2004067239-A1.
 FN
 XX 08-APR-2004.
 PD
 XX 02-OCT-2002; 2002US-00263153.
 PF
 XX 02-OCT-2002; 2002US-00263153.
 PR
 XX (MAIN/) MAINE G T.
 PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 XX
 XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 FI
 XX WPI; 2004-304563/28.
 DR P-PSDB; ADO23614.
 DR
 XX Novel purified polypeptide having sequence identity to amino acid
 PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX
 XX Example 2; Fig 25; 114pp; English.
 XX
 CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
 CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX
 XX Sequence 7322 BP; 1887 A; 1844 C; 1934 G; 1657 T; 0 U; 0 Other;
 SQ
 Query Match 81.5%; Score 2690.4; DB 12; Length 7322;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGTGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGGAGAGA 60
 DB 1 CCGACACCATCGAATGTGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGGAGAGA 60
 QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACTATACGATGTGCGAGAGTATGCCG 120
 DB 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACTATACGATGTGCGAGAGTATGCCG 120
 QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCAGCTTTCGCGAAAA 180
 DB 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCAGCTTTCGCGAAAA 180
 QY 181 CCGGGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTCCCAACCGGTGGCAC 240
 DB 181 CCGGGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTCCCAACCGGTGGCAC 240

QY 241 AACAACTGGCGGCGCAACAGTCGTTGCTGATTGSCGTTGCCACCTCCAGTCTGGCCCTGC 300
 DB 241 AACAACTGGCGGCGCAACAGTCGTTGCTGATTGSCGTTGCCACCTCCAGTCTGGCCCTGC 300
 QY 301 ACGCGCGTCGCAAAATTTGTGCGCGCGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
 DB 301 ACGCGCGTCGCAAAATTTGTGCGCGCGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
 QY 361 TGGTGGTGTGATGGTAGAACGAAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC 420
 DB 361 TGGTGGTGTGATGGTAGAACGAAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC 420
 QY 421 TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAACCTATCCGCTGGATGACCAAGATGCCA 480
 DB 421 TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAACCTATCCGCTGGATGACCAAGATGCCA 480
 QY 481 TTGCTGTGGAAGCTGCCTGCACCTAAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
 DB 481 TTGCTGTGGAAGCTGCCTGCACCTAAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
 QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAACCGGTACGCGACTGGGCGTGGAGATC 600
 DB 541 CACCCATCAACAGTATTAATTTCTCCCATGAACCGGTACGCGACTGGGCGTGGAGATC 600
 QY 601 TGGTTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
 DB 601 TGGTTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
 QY 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
 DB 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
 QY 721 CGGAAACGGGAAGCGGACTGAGTGCATGTCGGTTCGGTTTCAACAAACCATCAAAATGCTGA 780
 DB 721 CGGAAACGGGAAGCGGACTGAGTGCATGTCGGTTCGGTTTCAACAAACCATCAAAATGCTGA 780
 QY 781 ATGAGGCGCATCGTTCCACATCGCATGCTGTTGCCAACAGTACAGTGGCGCTGGGCGCAA 840
 DB 781 ATGAGGCGCATCGTTCCACATCGCATGCTGTTGCCAACAGTACAGTGGCGCTGGGCGCAA 840
 QY 841 TGGCGCGCAATTAACGAGTCCGGCTCGCGCTGCGGATATCTCGGTAGTGGGATACG 900
 DB 841 TGGCGCGCAATTAACGAGTCCGGCTCGCGCTGCGGATATCTCGGTAGTGGGATACG 900
 QY 901 ACCATACCGAAGACAGTCAATGTTATATCCCGCGGTTAACACCATCAAAACAGATTTTC 960
 DB 901 ACCATACCGAAGACAGTCAATGTTATATCCCGCGGTTAACACCATCAAAACAGATTTTC 960
 QY 961 GCCTGTGCGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGA 1020
 DB 961 GCCTGTGCGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGA 1020
 QY 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
 DB 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
 QY 1081 CGCAAAACCGCTCTCCCGCGCGTTGCGGATTCATTATGCGAGCTGGCACAGAGTTT 1140
 DB 1081 CGCAAAACCGCTCTCCCGCGCGTTGCGGATTCATTATGCGAGCTGGCACAGAGTTT 1140
 QY 1141 CCCGACTGGAAACGGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 1200
 DB 1141 CCCGACTGGAAACGGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 1200
 QY 1201 GCACAATTCATGTTTGAAGCTTATCATCGACGGTGCACCAATGCTTCTGGCG 1260
 DB 1201 GCACAATTCATGTTTGAAGCTTATCATCGACGGTGCACCAATGCTTCTGGCG 1260
 QY 1261 TCAGGCGAGCCATCGGAAGCTGTGATGCTGTGCGAGTGTGTAATCCTGCAATATTCG 1320
 DB 1261 TCAGGCGAGCCATCGGAAGCTGTGATGCTGTGCGAGTGTGTAATCCTGCAATATTCG 1320

QY 1321 TGTGCTCAAGCGGCACCTCCGTTCTGGATATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCGTTCTGGATATATGTTTTTGGCGGACATCATACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATAATGTTGGA 1440
Db 1381 CTGGCAAAATATTCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTGAGGGGATACAATTTTACACAGGAAACAGCCAGTCGGTTTAGGTFTTTCAGCA 1500
Db 1441 ATTGTGAGGGGATACAATTTTACACAGGAAACAGCCAGTCGGTTTAGGTFTTTCAGCA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACCGGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACCGGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAAATCGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCGCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCGCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGACACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGACACCCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAACAAACACATGAATGCAGACACCATTTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAACAAACACATGAATGCAGACACCATTTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAGATTAATGTTGATGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGATTAATGTTGATGCGGACCTTC 2280
QY 2281 AAGGTCACACATCCAAACCGTTCTGCGCTGAGCGGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACATCCAAACCGTTCTGCGCTGAGCGGAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTTCGAAAATCTCTGCTGACATGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTTCGAAAATCTCTGCTGACATGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTTACGAGGAAG 2460

Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTTGAAGTCTTTACGAGGAAG 2460
QY 2461 TTGCGGAAAGATCCACGTTATTGCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTTATTGCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGAGATGTCGCGTTTCTGGTATGCGCGTCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGCGTTTCTGGTATGCGCGTCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAACAATAACAACAACTCGGGATCGAGGGGAAGATTTCAGAAATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACAACTCGGGATCGAGGGGAAGATTTCAGAAATTC 2700
RESULT 10
ADO23603
ID ADO23603 standard; DNA; 7352 BP.
XX
AC ADO23603;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-ToxoP30del14C(52-294aa) fusion protein.
XX
KW P30 antigen; Toxo30del13C; Toxo30del2C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
DR WPI; 2004-304563/28.
DR P-PSDB; ADO23604.
XX
PT Novel purified polypeptide having sequence identity to amino acid sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C, Toxo30del2C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG antibodies to Toxoplasma gondii.
XX
PS Example 2; Fig 19; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully defined Toxo30del13C, Toxo30del2C and ToxoP30 MIX1 sequences, having 1-6 amino acids added to the C-terminus of the amino acid sequence of Toxo30del2C P30 antigen sequence, or comprising an amino acid sequence in which at least one of the five C-terminal cysteine amino acids of the amino acid sequence of Toxo30del13C P30 antigen sequence is substituted with alanine, or comprising the amino acid sequence chosen from MBP-Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1 fusion proteins. (I) is useful for detecting the presence of IgM antibodies to Toxoplasma gondii in a test sample, which involves contacting the test sample suspected of containing the IgM antibodies

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.

XX Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	1	CCGACACCATCGAATGGTGCRAAACCTTTCGGGTATGGCATGATAGGCCCGGAAGAGA	60
Db	1	CCGACACCATCGAATGGTGCRAAACCTTTCGGGTATGGCATGATAGGCCCGGAAGAGA	60
Qy	61	GTCAATTGAGGGTGGTGAATGTGAATCGTAAACAGTAAAGTATACGATGTGCGCAGAGTATGCCG	120
Db	61	GTCAATTGAGGGTGGTGAATGTGAATCGTAAACAGTAAAGTATACGATGTGCGCAGAGTATGCCG	120
Qy	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCCACGTTTCTCGGAAAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCCACGTTTCTCGGAAAA	180
Qy	181	CGCGGAAAAAGTGGAGCGGCGATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC	240
Db	181	CGCGGAAAAAGTGGAGCGGCGATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC	240
Qy	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Qy	301	ACGCGCGTGCAGAAATGTGCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCGTGCAGAAATGTGCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Qy	361	TGTTGTTGTCGATGGTGAAGCGGCGTGAAGCGCTGTAAGCGCGCGTGAACATC	420
Db	361	TGTTGTTGTCGATGGTGAAGCGGCGTGAAGCGCTGTAAGCGCGCGTGAACATC	420
Qy	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACAGGATGCCA	480
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACAGGATGCCA	480
Qy	481	TTGCTGTGGAAGCTGCTGACATAATGTTCGGCGTTATTTCTGATGTCTCTGACAGA	540
Db	481	TTGCTGTGGAAGCTGCTGACATAATGTTCGGCGTTATTTCTGATGTCTCTGACAGA	540
Qy	541	CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
Qy	601	TGTFCCGATTGGGTACACAGCAATCGCTGTTAGCGGGCCCAATTAAGTTCTGCTCGG	660
Db	601	TGTFCCGATTGGGTACACAGCAATCGCTGTTAGCGGGCCCAATTAAGTTCTGCTCGG	660
Qy	661	CGGCTCTGCTCTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGGCTCTGCTCTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAAACGGGAAGGCGACTGGAGTGCATGTCGGTTCGCAACCAACCATCAATGCTGA	780
Db	721	CGGAAACGGGAAGGCGACTGGAGTGCATGTCGGTTCGCAACCAACCATCAATGCTGA	780
Qy	781	ATGAGGCGCATGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA	840
Db	781	ATGAGGCGCATGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA	840
Qy	841	TGCGCGCATTCAGGATCGCGGTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTCAGGATCGCGGTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACATCAACAGGATTTTC	960
pb	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACATCAACAGGATTTTC	960

Qy	961	GCCTGCTGGGGCAAAACAGCGTGCAGCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAACAGCGTGCAGCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Qy	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAAAAGAAAAACACCTCGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAAAAGAAAAACACCTCGCGCCCAATA	1080
Qy	1081	CGMAACCGCTCTCCCGCGGTGGCCGATTCATTAATGACAGCTGGCAGCAGGTTT	1140
Db	1081	CGMAACCGCTCTCCCGCGGTGGCCGATTCATTAATGACAGCTGGCAGCAGGTTT	1140
Qy	1141	CCGACTGGAAGCGGCGCAGTGCAGCAACGCAATTAATGAGTTAGCTCACTCATTAG	1200
Db	1141	CCGACTGGAAGCGGCGCAGTGCAGCAACGCAATTAATGAGTTAGCTCACTCATTAG	1200
Qy	1201	GCAAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCAAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260
Qy	1261	TCAGGACCCATCGGAAGCTGTGATGGCTGTGCGAGGTGTAATCACTGCATAATTCG	1320
Db	1261	TCAGGACCCATCGGAAGCTGTGATGGCTGTGCGAGGTGTAATCACTGCATAATTCG	1320
Qy	1321	TGTCGCTCAAGGCGCACTCCCGTCTGGATAATGTTTTTTCGCGCGACATCAACCGTT	1380
Db	1321	TGTCGCTCAAGGCGCACTCCCGTCTGGATAATGTTTTTTCGCGCGACATCAACCGTT	1380
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Qy	1441	ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500
Qy	1501	GCACTTTCACCAACAGGACCATAGATTATGAAATCTGAAAGGTAACTGGTAACTCGG	1560
Db	1501	GCACTTTCACCAACAGGACCATAGATTATGAAATCTGAAAGGTAACTGGTAACTCGG	1560
Qy	1561	ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCCGTTAAGAAATTCGAGAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCCGTTAAGAAATTCGAGAAGAT	1620
Qy	1621	ACCGGAATTAAGTACCGTTCGGAATAACTGGGAAGAAATTCACACAGGTT	1680
Db	1621	ACCGGAATTAAGTACCGTTCGGAATAACTGGGAAGAAATTCACACAGGTT	1680
Qy	1681	CGGCAACTGGCGATGGCCCTGACATTTCTCTGGGCACACACCGCTTGGTGGCTAC	1740
Db	1681	CGGCAACTGGCGATGGCCCTGACATTTCTCTGGGCACACACCGCTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCGTGTGGCTGAAATCACCCCGAATAAGCGTTCCAGGACAAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCGTGTGGCTGAAATCACCCCGAATAAGCGTTCCAGGACAAAGCTGTAT	1800
Qy	1801	CGGTTTACCTGGGATCGGTTACAGCGCAAGCTGATTCCTTACCAGATCGCTGTT	1860
Db	1801	CGGTTTACCTGGGATCGGTTACAGCGCAAGCTGATTCCTTACCAGATCGCTGTT	1860
Qy	1861	GAGCGGTTATCGCTGATTTTAAACAAAGATCTCTGCCGAACCCCGCAAAAACTGGGAA	1920
Db	1861	GAGCGGTTATCGCTGATTTTAAACAAAGATCTCTGCCGAACCCCGCAAAAACTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGCGGTAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGCGGTAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACCTTACCTGGCCGCTGATTTGCTGCTGACGGGGGTTATGCTTCAAG	2040
Db	1981	CTGCAAGAACCGTACCTTACCTGGCCGCTGATTTGCTGCTGACGGGGGTTATGCTTCAAG	2040

```
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAACGCGCGGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAACGCGCGGTGG 2220
QY 2221 GCATGCTCAACATCGACACACGAGCAAGTGAATATGCTGTAAACGCTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACACGAGCAAGTGAATATGCTGTAAACGCTACTGCGGACCTTC 2280
QY 2281 AAGGCTCAACCATCAACACCGTTCGTCGCGTCTGAGCGAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGCTCAACCATCAACACCGTTCGTCGCGTCTGAGCGAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTATAAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTATAAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAAGAGCGGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAAGAGCGGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACACACAAATAACAATAACAACACCTCGGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACACACACAAATAACAATAACAACACCTCGGGATCGAGGAGGATTCAGAAATTC 2700

RESULT 11
ADO23598
ID ADO23598 standard; DNA; 7370 BP.
AC
XX
XX
XX
DT 01-JUL-2004 (first entry)
DE DNA encoding MBP-ToxoP30del13C(52-300aa) fusion protein.
KW
XX
KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; db; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
FN US2004067239-A1.
PD
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN//) MAINE G T.
PA (PATE//) PATEL C B.
PA (GINS//) GINSBURG S R.
```

```
PA (BLIE//) BLIESE T R.
XX
XX Maine GT, Patel CB, Gineburg SR, Bliese TR;
XX
DR WPI; 2004-304563/28.
DR P-PSDB; ADO23599.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
PS Example 2; Fig 15; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1901 A; 1855 C; 1949 G; 1665 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGCGCATGTAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGCGCATGTAGCGCCGGAAGAGA 60
QY 61 GTCATTTCCAGGTGTGTAATGTGAAACCAAGTAACTTATAGATGTCGACAGATATGCG 120
Db 61 GTCATTTCCAGGTGTGTAATGTGAAACCAAGTAACTTATAGATGTCGACAGATATGCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTAACAGCCAGCCAGCCAGCTTTCTCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTAACAGCCAGCCAGCCAGCTTTCTCGAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTAATTCCTCCAAACCGCTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTAATTCCTCCAAACCGCTGGCAC 240
QY 241 AACAACTCGCGGCAACACAGTTCGTTGCTGATTGGGTTTGGCCACTTCCAGTCTGGCCCTGC 300
Db 241 AACAACTCGCGGCAACACAGTTCGTTGCTGATTGGGTTTGGCCACTTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTCGTCGGCGATTAATCTCGGCCGATCAACTGGGTGCCAGC 360
Db 301 ACGCGCGTTCGCAAAATTCGTCGGCGATTAATCTCGGCCGATCAACTGGGTGCCAGC 360
QY 361 TGGTGGTGTGATGTAGAACGAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAACATC 420
Db 361 TGGTGGTGTGATGTAGAACGAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAACATC 420
QY 421 TTCTCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
```

Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTCCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db 601 TGGTCCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
Qy 661 CGCGTCTCGCTGCGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGACATGGAGTGCATGTCGGTTTCAACAAACATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACATGGAGTGCATGTCGGTTTCAACAAACATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGGCAA 840
Qy 841 TGGCGGCCATTACCGAGTCCGGGCTGCGGCTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGCGGCTGGTGGCGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCACCATCAAAACAGATTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCACCATCAAAACAGATTTC 960
Qy 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGTCCTCTCACTGTTGAAAGAAACCAACCTGCGGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGTCCTCTCACTGTTGAAAGAAACCAACCTGCGGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGTTGCGGATTCATTAATGCAGCTGGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGTTGCGGATTCATTAATGCAGCTGGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTAG 1200
Qy 1201 GCACAAATCTCATGTTGACAGCTTATCATGCTGACGCGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTGACAGCTTATCATGCTGACGCGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGTTATGCTGTGCTGCTGAGTCACTGCAATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGTTATGCTGTGCTGCTGAGTCACTGCAATAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAAATGTTTTTGGCGGACATCAAAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAAATGTTTTTGGCGGACATCAAAACGGTT 1380
Qy 1381 CTGGCAATATCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAATATCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
Qy 1501 GCATTCACCAACAGGACCATGATTAATGAAACTGAAAGAGTAAATGTTGTAATCTCG 1560
Db 1501 GCATTCACCAACAGGACCATGATTAATGAAACTGAAAGAGTAAATGTTGTAATCTCG 1560
Qy 1561 ATTAACGGGATTAAGGCTTAACGGTCTCGCTGAGTCCGGTAAAGTAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGGATTAAGGCTTAACGGTCTCGCTGAGTCCGGTAAAGTAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT 1680

RESULT 12

AD023639

ID AD023639 standard; DNA; 7370 BP.

Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACACGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACACGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGGTTTCCAGGACAACTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGGTTTCCAGGACAACTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCCGTACGTTTACAACGGCAAGCTGATTGTTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCCGTACGTTTACAACGGCAAGCTGATTGTTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Qy 1981 CTGCAAGAACCGTATCTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTATCTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAACAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAACTGCTTTTAAAGGCGAAACAGCGATGACCATTAACGCGCCGTGG 2220
Db 2161 TCCATCGCAGAACTGCTTTTAAAGGCGAAACAGCGATGACCATTAACGCGCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACCGCAAGTGAATTAATGTTGTTAAAGGCGTAACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCGCAAGTGAATTAATGTTGTTAAAGGCGTAACTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAAACGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAAACGCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGATTCTCGAATACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGATTCTCGAATACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGGTTTCTGGTATGCCGTGGTACTCGCGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGTTTCTGGTATGCCGTGGTACTCGCGTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACCAACAAACAAATTAACAACTTCCGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATTAACAACTTCCGATCGAGGGAAGGATTTTCAGAAATTC 2700

XX AC ADO23639;
XX DT 01-JUL-2004 (first entry)
XX DE DNA encoding MBP-ToxoP30MIX1 fusion protein.
XX KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX OS Toxoplasma gondii.
OS Synthetic.
XX PN US2004067239-A1.
XX PD 08-APR-2004.
XX PF 02-OCT-2002; 2002US-00263153.
XX PR 02-OCT-2002; 2002US-00263153.
XX PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIESE/) BLIESE T R.
XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX DR WPI; 2004-304563/28.
DR P-PSDB; ADO23640.
XX PT Novel purified polypeptide having sequence identity to amino acid
PT antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of Igm or Igg
PT antibodies to Toxoplasma gondii.
XX Example 5; Fig 31; 114pp; English.
XX The invention relates to a purified p30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C p30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C p30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of Igm
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the Igm antibodies
CC with a composition comprising (I) and detecting the presence of (I)/Igm
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX SQ Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;
Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAAATTCAGGTGGTGAATGTGAACACGTAACGTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAAATTCAGGTGGTGAATGTGAACACGTAACGTTATACGATGTCGAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGGTTTCCTCCGGTGGTGAACACGAGCCAGCCAGCTTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGGTTTCCTCCGGTGGTGAACACGAGCCAGCCAGCTTTCTCGGAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGCATGGCGGAGCTGAATTTACATTCCTCCCAACCGCTGGCAC 240

Db 181 CGCGGAAAAAGTGAAGCGCGCATGGCGGAGCTGAATTTACATTCCTCCCAACCGCTGGCAC 240
Qy 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTTCGCAAAATTTGTCGCGCGGATTAATCTCTCGCGCGGATCAATCTGGGTGGCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTTGTCGCGCGGATTAATCTCTCGCGCGGATCAATCTGGGTGGCAGCG 360
Qy 361 TGGTGGTGTCTGATGTAAGACGAGCGCGTTCGAGCGCTGTAAAGCGCGGTGACACAATC 420
Db 361 TGGTGGTGTCTGATGTAAGACGAGCGCGTTCGAGCGCTGTAAAGCGCGGTGACACAATC 420
Qy 421 TTCTCGCGCAACCGGTTCAGTGGGCTGATCACTAATCTATCCCTGGATCACCAGATGCCA 480
Db 421 TTCTCGCGCAACCGGTTCAGTGGGCTGATCACTAATCTATCCCTGGATCACCAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGTTAATTTCTTGATGTCTCTGACCA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGTTAATTTCTTGATGTCTCTGACCA 540
Qy 541 CACCATCAACAGTATTATTTCTCCCATGAAGACGTTACCGACTGGCGGTGGAGCATC 600
Db 541 CACCATCAACAGTATTATTTCTCCCATGAAGACGTTACCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGGTTCACAGCAAAATCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGGCATTTGGGTTCACAGCAAAATCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCACAAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCACAAACCAATGCAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGGCGTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGGCGTGGCGGCA 840
Qy 841 TGCAGCCATTAACGAGTCCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGCAGCCATTAACGAGTCCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
Qy 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGCTGA 1020
Db 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGCTGA 1020
Qy 1021 AGGGCAATCAGCTGTGTTGCCCGTCTCACTGTGTAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTTGCCCGTCTCACTGTGTAAGAAAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATGCACTGCAAGCGGTGCACCAATGTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATGCACTGCAAGCGGTGCACCAATGTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGTTAGTGTGAGGTGCTGAATTCATCTGCATTAATTCG 1320

Db 1261 TCAGGACGATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAATTTCG 1320
Qy 1321 TGTCCCTCAAGCGCACATCCCGTCTCGGATATGTTTTTGGCCGACATATAACGGTT 1380
Db 1321 TGTCCCTCAAGCGCACATCCCGTCTCGGATATGTTTTTGGCCGACATATAACGGTT 1380
Qy 1381 CTGGCAAAATATTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGGCAAAATATTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Qy 1441 ATTGTAGCGGATACAAATTTTACACAGGAAACAGCCAGTCGGTTAGGTGTTTTTCAGGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTACACAGGAAACAGCCAGTCGGTTAGGTGTTTTTCAGGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATATTATGAACCTGAAGAGGTAAACTGCTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATATTATGAACCTGAAGAGGTAAACTGCTAATCTGG 1560
Qy 1561 ATTAAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTTAAGTACACGTTGAGCATCCGATCCGATAACTGGGAAGAGAAATTCGCCACAGGTT 1680
Db 1621 ACCGGAATTTAAGTACACGTTGAGCATCCGATCCGATAACTGGGAAGAGAAATTCGCCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTGGCTGAATCAATCAACCCGACAAAGCGTTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAATCAATCAACCCGACAAAGCGTTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTCGGATGCGTACGTTACAAACGCAAGCTGATGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTCGGATGCGTACGTTACAAACGCAAGCTGATGCTTACCGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGCTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGCTGGATACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGAATGAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGTCTGCTTTAATAAAGCGGAAACACGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGTCTGCTTTAATAAAGCGGAAACACGCGATGACCATCAACGGCCGCTGG 2220
Qy 2221 GCATGTCCAAATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGTCCAAATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGTCACCAATCCAAACCTGTTGGTGGCTGCTGAGCCAGGTATTAAACGGCCGAGT 2340
Db 2281 AAGGTCACCAATCCAAACCTGTTGGTGGCTGCTGAGCCAGGTATTAAACGGCCGAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGATGAAGGTCTG 2400

Qy 2401 GAAGCGGTTAATAAGACAAACCCCTGGTGCCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCCCTGGTGCCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCACGCTATTGGCCGACCATGGAACCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATTGGCCGACCATGGAACCCCAAGAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGGTATGCCGTGGGTACTGCGGTGATCAAGGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGGTATGCCGTGGGTACTGCGGTGATCAAGGCC 2580
Qy 2581 GCCAGCGTGTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCCAGCGTGTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Qy 2641 AACCAACAAACAATAAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACCAACAAACAATAAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

RESULT 13
ADO23649
ID ADO23649 standard; DNA; 7370 BP.
XX ADO23649;
AC ADO23649;
XX ADO23649;
DT 01-JUL-2004 (first entry)
XX DNA encoding MBP-Toxop30MIX5 fusion protein.
XX P30 antigen; Toxo30del3C; Toxo30del12C; Toxop30 MIX1;
KW MBP-Toxo30del3C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX Toxoplasma gondii.
OS Synthetic.
XX US2004067239-A1.
XX 08-APR-2004.
XX 02-OCT-2002; 2002US-00263153.
XX 02-OCT-2002; 2002US-00263153.
XX (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23650.
XX Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del3C,
PT Toxo30del12C, Toxo30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX Example 5; Fig 35; 114pp; English.
XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del3C, Toxo30del12C and Toxo30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del3C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del3C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (i) is useful for detecting the presence of IgM

CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (i) and detecting the presence of (i)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CCGACACCATCGATGCTGCAAAACCTTTCGGGTATGCGCATATAGCGCCCGAGAGA	60
DB	1	CGACACCATCGAATGGTGCAAAACCTTTCGGGTATGCGCATATAGCGCCCGAGAGA	60
QY	61	GTCATTCAGGGTGGTGAATGTGAACACAGTAACCGTATACGATGTCGACAGTATGCCG	120
DB	61	GTCATTCAGGGTGGTGAATGTGAACACAGTAACCGTATACGATGTCGACAGTATGCCG	120
QY	121	GTGTCCTTTATCAGACCGTTCCTCCGCGTGGTGAACACGCGCACCGTTCCTGCGAAAA	180
DB	121	GTGTCCTTTATCAGACCGTTCCTCCGCGTGGTGAACACGCGCACCGTTCCTGCGAAAA	180
QY	181	CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAACATTCGCCAACCGCGTGGCAC	240
DB	181	CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAACATTCGCCAACCGCGTGGCAC	240
QY	241	AACAACTCGCGGCAACAGTCTGCTGATTTGGGGTTGCCACCTTCAGTCTGGGCCCTGC	300
DB	241	AACAACTCGCGGCAACAGTCTGCTGATTTGGGGTTGCCACCTTCAGTCTGGGCCCTGC	300
QY	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGC	360
DB	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGC	360
QY	361	TGGTGGTTCGATGTAGACAAGCGCGGTGAAAGCTGTAAGCGCGGTGCAATATC	420
DB	361	TGGTGGTTCGATGTAGACAAGCGCGGTGAAAGCTGTAAGCGCGGTGCAATATC	420
QY	421	TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAG	480
DB	421	TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAG	480
QY	481	TTGCTGTGGAAGCTGCCTGCAATAATGTTCCGGCGTATTTCTGATGTTCTCTGACCAGA	540
DB	481	TTGCTGTGGAAGCTGCCTGCAATAATGTTCCGGCGTATTTCTGATGTTCTCTGACCAGA	540
QY	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGGCATGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGGCATGGGCGTGGAGCATC	600
QY	601	TGGTCGCAATGGGTCCACAGCAAACTCGGCTGTAGCGGGCCCATTAAGTCTGCTCGG	660
DB	601	TGGTCGCAATGGGTCCACAGCAAACTCGGCTGTAGCGGGCCCATTAAGTCTGCTCGG	660
QY	661	CGCGTCTCGGTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTCGGTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAAACGGGAAGCGAGTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAAACGGGAAGCGAGTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTCCCACTCGCATGCTGGTTGGTCCAAACGATCAGATGGCGCTGGCGCAA	840
DB	781	ATGAGGGCATCGTCCCACTCGCATGCTGGTTGGTCCAAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGGCGCATTAACAGTCCGGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATAGC	900
DB	841	TGCGGCGCATTAACAGTCCGGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATAGC	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACATCAAAAGGATTTTC	960

DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
DB	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGGCCCTCTCACTGTTGAAAGAAAACCAACCTCGGCGCCCAATA	1080
DB	1021	AGGGCAATCAGCTGTTGGCCCTCTCACTGTTGAAAGAAAACCAACCTCGGCGCCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGCGACAGAGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGCGACAGAGTTT	1140
QY	1141	CCGACATGGAAGGGGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG	1200
DB	1141	CCGACATGGAAGGGGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGAGCGTGCACCAATGCTTCGGG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGAGCGTGCACCAATGCTTCGGG	1260
QY	1261	TCAGGACGCCATCGAAGCTGTGGTATGCGTGTGAGGTCTGTAATCACTGCAATAATTCG	1320
DB	1261	TCAGGACGCCATCGAAGCTGTGGTATGCGTGTGAGGTCTGTAATCACTGCAATAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCCGTCTTGGATTAATGTTTTTGGCGGACATCAATACGGTT	1380
DB	1321	TGTCGCTCAAGCGCACCTCCCGTCTTGGATTAATGTTTTTGGCGGACATCAATACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
QY	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACAGCGAGTCCGTTTAGGTGTTTTCAG	1500
DB	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACAGCGAGTCCGTTTAGGTGTTTTCAG	1500
QY	1501	GCATTTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCTGTAATCTGG	1560
DB	1501	GCATTTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCTGTAATCTGG	1560
QY	1561	ATTTAAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTTAAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTTAAAGTCACCGTTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCACAGGT	1680
DB	1621	ACCGGAATTTAAAGTCACCGTTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCACAGGT	1680
QY	1681	CGGCAACTGGCGATGGCCCTGACATTATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
DB	1681	CGGCAACTGGCGATGGCCCTGACATTATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGGCAAGAGCGTTCCAGACAAGCTGTAT	1800
DB	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGGCAAGAGCGTTCCAGACAAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCGCTGTT	1860
DB	1801	CGGTTTACCTGGGATGGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCGCTGTT	1860
QY	1861	GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCGCAACCCGCCCAAAACCTTGGGAA	1920
DB	1861	GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCGCAACCCGCCCAAAACCTTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTACGGGGGTTATCGGTTCAAG	2040

QY 541 CACCCATCAACAGTATTATTTCTCCATGAGACGGTACCGACTGGCGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCATGAGACGGTACCGACTGGCGCTGGAGCATC 600
QY 601 TGGTCGATTTGGGTCAACAGAAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
Db 601 TGGTCGATTTGGGTCAACAGAAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTCGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTCGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACATGGAGTGCATGTCGGTTTTCAACAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGACATGGAGTGCATGTCGGTTTTCAACAACCAATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGAA 840
QY 841 TGCGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATACG 900
Db 841 TGCGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCAATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCAATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGACCGCTTGCTGCAACTCTCAAGGCGAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGACCGCTTGCTGCAACTCTCAAGGCGAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGCAACAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGCAACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATGACTGCACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATGACTGCACGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTTTTCACGA 1500
QY 1501 GCACCTTACCACAGGACCATAGATATGAAAATGGAAGGTTAACTGTAATCTCG 1560
Db 1501 GCACCTTACCACAGGACCATAGATATGAAAATGGAAGGTTAACTGTAATCTCG 1560
QY 1561 ATTTAAGCGGATAAAGGCTTAAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTTAAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGCAACTGCGCATGCGCCCTGACATTATCTTTCTGGGCACACGACCGCTTTGGTGCTAC 1740
Db 1681 GCGCAACTGCGCATGCGCCCTGACATTATCTTTCTGGGCACACGACCGCTTTGGTGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGACAAAGGTTCCAGGAACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGACAAAGGTTCCAGGAACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAACGGAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAACGGAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACCTGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAAGACGCTGGGCGTGGATTAACGCTGGCGGAAACG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAAGACGCTGGGCGTGGATTAACGCTGGCGGAAACG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAAAAAACAACATGAATGACAGACCCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAAAAAACAACATGAATGACAGACCCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCATCAACCGTTGCTGGCGTGTGAGCGCAGGATTAACCGCCCGCAGT 2340
Db 2281 AAGGTCACCATCAACCGTTGCTGGCGTGTGAGCGCAGGATTAACCGCCCGCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAATCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAATCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCAAAAGATTCACGATTGTCGCCCAACCATGGAAGAAACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCAAAAGATTCACGATTGTCGCCCAACCATGGAAGAAACGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCACGGTCTGTCAGATGTCGATGAAGCCCTGAAAGACCGCCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCACGGTCTGTCAGATGTCGATGAAGCCCTGAAAGACCGCCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACAACAACAACAACAACAACAACAACAACAACAACA 2700
Db 2641 AACAAACAACAATAACAATAACAACAACAACAACAACAACAACAACAACAACAACAACA 2700

ADO23594
 ID ADO23594 standard; DNA; 7403 BP.
 XX AC ADO23594;
 XX DT 01-JUL-2004 (first entry)
 XX DE DNA encoding MBP-ToxoP30del12(52-311aa) fusion protein.
 XX KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
 KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; db; gene.
 XX OS Toxoplasma gondii.
 OS Synthetic.
 XX PN US2004067239-A1.
 XX PD 08-APR-2004.
 XX PF 02-OCT-2002; 2002US-00263153.
 XX PR 02-OCT-2002; 2002US-00263153.
 XX PA (MAIN/) MAINE G T.
 PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 DR WPI; 2004-304563/28.
 DR P-PSDB; ADO23595.
 XX PT Novel purified polypeptide having sequence identity to amino acid
 PT sequence of p30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX PS Example 2; Fig 11; 114pp; English.
 XX CC The invention relates to a purified p30 antigen (I) chosen from 3 fully
 CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX SQ Sequence 7403 BP; 1907 A; 1863 C; 1961 G; 1672 T; 0 U; 0 Other;
 Query Match 81.5%; Score 2690.4; DB 12; Length 7403;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGGTGAACACCTTCGCGGTATGGCATATAGCCCGCGGAAGAGA 60
 DB 1 CCGACACCATCGAATGGTGAACACCTTCGCGGTATGGCATATAGCCCGCGGAAGAGA 60
 QY 61 GTCAATTCAGGGTGGTGAATGTGAACACCTTCAGATATACGATGTGCGAGATATGCCG 120
 DB 61 GTCAATTCAGGGTGGTGAATGTGAACACCTTCAGATATACGATGTGCGAGATATGCCG 120
 QY 121 GTCTCTTATCAGACCGTTTCGCGGTGTGTGAACACCGACCGACCGTTCGCGAAA 180
 DB 121 GTCTCTTATCAGACCGTTTCGCGGTGTGTGAACACCGACCGACCGTTCGCGAAA 180

QY 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATCCCAACCGCTGGCAC 240
 DB CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATCCCAACCGCTGGCAC 240
 QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
 DB AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
 QY 301 ACGCCGCTCGCAAAATTTGTGCGCGGATTAATCTCGCGCGGATCAACTGGTGGCGACGC 360
 DB ACGCCGCTCGCAAAATTTGTGCGCGGATTAATCTCGCGCGGATCAACTGGTGGCGACGC 360
 QY 361 TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCGCTGTAAGCGCGGCTGCACAATC 420
 DB TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCGCTGTAAGCGCGGCTGCACAATC 420
 QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCACTATCCGCTGGATGACAGATGCCA 480
 DB TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCACTATCCGCTGGATGACAGATGCCA 480
 QY 481 TTGCTGTGAAGCTGCCTGCACATAATGTTCCGGCGTTTATTTCTTGATGTCTCTGACCAGA 540
 DB TTGCTGTGAAGCTGCCTGCACATAATGTTCCGGCGTTTATTTCTTGATGTCTCTGACCAGA 540
 QY 541 CACCCATCAACAGTATTTATTTCTCCATGAAGCGGTACGCACTGGGCGTGGAGCATC 600
 DB CACCCATCAACAGTATTTATTTCTCCATGAAGCGGTACGCACTGGGCGTGGAGCATC 600
 QY 601 TGGTTCGCAATTTGGCTCACAGCAATCGCGCTGTTAGCGGCGCCATTAAGTTCCTCGG 660
 DB TGGTTCGCAATTTGGCTCACAGCAATCGCGCTGTTAGCGGCGCCATTAAGTTCCTCGG 660
 QY 661 CGGCTCTGCTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATCAGCCCATAG 720
 DB CGGCTCTGCTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATCAGCCCATAG 720
 QY 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATCTGA 780
 DB CGGAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATCTGA 780
 QY 781 ATGAGGCAATCGTTCCACTTCGCGATGCTGTTGCCAACGATCAGATGGCGTGGGCGCAA 840
 DB ATGAGGCAATCGTTCCACTTCGCGATGCTGTTGCCAACGATCAGATGGCGTGGGCGCAA 840
 QY 841 TGGCGCCATTACCGAGTCCGGCTCGCGTGGTGGCGGATATCTCGTGGTGGGATACG 900
 DB TGGCGCCATTACCGAGTCCGGCTCGCGTGGTGGCGGATATCTCGTGGTGGGATACG 900
 QY 901 AGCATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
 DB AGCATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
 QY 961 GCTGCTGGGGCAAAACAGCGTGGACCGCTTCTGCACTCTCTCAGGGCCAGGGGGTGA 1020
 DB GCTGCTGGGGCAAAACAGCGTGGACCGCTTCTGCACTCTCTCAGGGCCAGGGGGTGA 1020
 QY 1021 AGGGCAATCAGCTGTTCGCCGTCTCACTGGTGAAGAAAAACCAACCCCTGGCGCCCAATA 1080
 DB AGGGCAATCAGCTGTTCGCCGTCTCACTGGTGAAGAAAAACCAACCCCTGGCGCCCAATA 1080
 QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
 DB CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
 QY 1141 CCGGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTATTAG 1200
 DB CCGGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTATTAG 1200
 QY 1201 GCACAAATTTCTATGTTTGACAGCTTATCATCGACTGACGCTGACCAATTCGTTCTGCGG 1260
 DB GCACAAATTTCTATGTTTGACAGCTTATCATCGACTGACGCTGACCAATTCGTTCTGCGG 1260

QY 1261 TCAGGACCCATCGAAGCTGTGGTATGCGCTGTGAGGTGTAATCACTCATTAATTCG 1320
Db 1261 TCAGGACCCATCGAAGCTGTGGTATGCGCTGTGAGGTGTAATCACTCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCCTGATTAATGTTTTTGGCGGACATCATTAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCCTGATTAATGTTTTTGGCGGACATCATTAACGGTT 1380
QY 1381 CTGGCAATATTTGAAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAATATTTGAAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTGTGAGGATTAACATTTTACACAGGAACACGAGTCCGTTAGGTGTTTTTCAGA 1500
Db 1441 ATTGTGAGGATTAACATTTTACACAGGAACACGAGTCCGTTAGGTGTTTTTCAGA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTATGAAACTGAAAGAGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGCATATGAATTCGAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCACTGGCGATGGCCCTGACATTTCTTGGGCACACAGCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCACTGGCGATGGCCCTGACATTTCTTGGGCACACAGCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCGTAGCTTACAAAGCGCAAGCTGATTGCTTACCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCGTAGCTTACAAAGCGCAAGCTGATTGCTTACCGATCCCTGTT 1860
QY 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTGCGAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTGCGAACCCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGAGCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGAGCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGGCTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGGCTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACAATGACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACAATGACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAACTCGACACGACCAAGTGAATTAATGTTAGCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACGACCAAGTGAATTAATGTTAGCGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTTGGCTGGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTTGGCTGGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400

Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTTATAAAGACAAACCCGCTGGGTCCCGTAGCGCTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTTTATAAAGACAAACCCGCTGGGTCCCGTAGCGCTGAAGTCTTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTTGCGGCCACATGGAAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTTGCGGCCACATGGAAAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700

Search completed: August 1, 2005, 04:28:37
Job time : 1101.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell updates/sec

Title: US-09-765-555B-15
Perfect score: 3300
Sequence: 1 ccgacacatcgaaatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_ats:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3300	100.0	3300	6	AX202426	Sequence
2	3300	100.0	3300	6	AX202427	Sequence
3	3226.4	97.8	3300	6	AX202430	Sequence
4	3215.2	97.4	3300	6	AX202429	Sequence
5	3207.2	97.2	3300	6	AX202428	Sequence
6	2705.8	82.0	7373	6	AX284144	Sequence
7	2705.8	82.0	8101	6	AX172306	Sequence
8	2701	81.8	6648	6	AX377531	Sequence
9	2700	81.8	7475	6	AX1965	Sequence
10	2700	81.8	7475	6	AR031992	Sequence
11	2700	81.8	7475	6	AR207294	Sequence
12	2698	81.8	9191	6	AX377532	Sequence
13	2676.2	81.1	6724	12	AF097412	Expressio
14	2636.8	79.9	6806	6	AX378208	Sequence
15	2631	79.7	6706	12	AF031088	Shuttle v
16	2598.2	78.7	6748	12	AF031813	Expressio
17	1609.4	48.8	5558	6	CQ794769	Sequence
18	1474	44.7	4700	6	CQ846980	Sequence
19	1462.4	44.3	4921	6	AX768201	Sequence

C	20	1462.4	44.3	4935	6	AX768200	Sequence
	21	1462.4	44.3	4945	6	AX768202	Sequence
	22	1462.4	44.3	4951	6	AX768203	Sequence
	23	1432.4	43.4	6755	12	AF140577	Integrati
	24	1417.6	43.0	1511	6	AX431576	Sequence
	25	1417.6	43.0	4297	12	AY219685	Expressio
	26	1417.6	43.0	4359	12	AY219686	Expressio
	27	1417.6	43.0	4419	12	AY219688	Expressio
	28	1417.6	43.0	4855	12	AY219689	Expressio
	29	1417.6	43.0	4968	12	AY219687	Expressio
	30	1417.6	43.0	6956	12	AY219682	Shuttle e
	31	1417.6	43.0	7018	12	AY219683	Shuttle e
	32	1417.6	43.0	7509	12	AY219684	Shuttle e
	33	1400.8	42.4	4566	12	AF177933	Cloning v
	34	1400.8	42.4	4635	12	AF177932	Cloning v
	35	1400.8	42.4	6399	12	AY302759	Expressio
	36	1400.8	42.4	6587	12	AY289542	Expressio
	37	1362	41.3	5201	6	AR493834	Sequence
	38	1362	41.3	5201	6	AR493835	Sequence
	39	1329.2	40.3	4806	12	AF050464	Expressio
C	40	1238.4	37.5	4803	12	AY243506	Cloning v
C	41	1238.4	37.5	5024	6	CQ846977	Sequence
	42	1237.2	37.5	1922	6	AX478528	Sequence
	43	1204	36.5	4506	12	AB116367	Expressio
	44	1204	36.5	4557	6	E13740	DNA sequenc
	45	1204	36.5	4557	6	AR373251	Sequence

ALIGNMENTS

RESULT 1
AX202426
LOCUS AX202426 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 14 from Patent WO0152620.
ACCESSION AX202426
VERSION AX202426.1 GI:15392173
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 14 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m1 and zinc finger protein ZFPm1"

ORIGIN	Query Match	100.0%;	Score 3300;	DB 6;	Length 3300;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3300;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	CGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGCGCATGATGATGCGCCCGGAAGAGA	60	
	1	CGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGCGCATGATGATGCGCCCGGAAGAGA	60	
Db	61	GTCAATTCAGGTGGTGAAT	GTGAACAGTAAACGTTATACGATGTCCGAGATGCGG	120	
QY	61	GTCAATTCAGGTGGTGAAT	GTGAACAGTAAACGTTATACGATGTCCGAGATGCGG	120	
Db	61	GTCAATTCAGGTGGTGAAT	GTGAACAGTAAACGTTATACGATGTCCGAGATGCGG	120	
QY	121	GTGTCCTCTTATCAGAC	CCGTCCTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTCTCGGAAAA	180	
Db	121	GTGTCCTCTTATCAGAC	CCGTCCTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTCTCGGAAAA	180	
QY	181	CGCGGAAAAAGTGGAA	CGCGGATGGCGGAGTATCCCAACCGCGTGGCAC	240	

Db 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTACATCCCAACCGCTGGCAC 240
Qy 241 ACAAACCTGGCGGCAAAACAGTCTGTTGCTGATGAGGCGTTGGCCACTCCAGTCTGGCCCTGC 300
Db 241 ACAAACCTGGCGGCAAAACAGTCTGTTGCTGATGAGGCGTTGGCCACTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGCTCGCAAAATTTGTCGGGCGGATTAATCTCGCGCGGATCAATCTGGTGGCCAGG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGGGCGGATTAATCTCGCGCGGATCAATCTGGTGGCCAGG 360
Qy 361 TGGTGGTGTGATGTGAAGCAAGCGCGCTGCGAAGCGCTGTAAGCGCGGCTGCAACAATC 420
Db 361 TGGTGGTGTGATGTGAAGCAAGCGCGCTGCGAAGCGCTGTAAGCGCGGCTGCAACAATC 420
Qy 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTAATTTCTTGATGTTCTCGACAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTAATTTCTTGATGTTCTCGACAGA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGACTGGCGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGACTGGCGCTGGAGCATC 600
Qy 601 TGGTGCATTCGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAATCTCTCTCGG 660
Db 601 TGGTGCATTCGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAATCTCTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCATCAAAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCATCAAAACCAATGCAATGCTGA 780
Qy 781 ATGAGGGCATCTGTTCCCACTCCGATGCTGTTGTCACAGATCAGATCGCGCTGGCGCAA 840
Db 781 ATGAGGGCATCTGTTCCCACTCCGATGCTGTTGTCACAGATCAGATCGCGCTGGCGCAA 840
Qy 841 TGCGCGCATTAACCGAGTCCGGCTGCGGCTGTTGTCGCGATATCTCGTGTAGTGGATACG 900
Db 841 TGCGCGCATTAACCGAGTCCGGCTGCGGCTGTTGTCGCGATATCTCGTGTAGTGGATACG 900
Qy 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCATCAACAGATTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCATCAACAGATTTTC 960
Qy 961 GCCTGTGGGCAAAACAGCGCTGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db 961 GCCTGTGGGCAAAACAGCGCTGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGCTGGCGGATTAATGAGTGTAGCTCACTCATTTAG 1140
Db 1081 CGCAACCGCTCTCCCGCGGCTGGCGGATTAATGAGTGTAGCTCACTCATTTAG 1140
Qy 1141 CCCGACTGGAAGCGGAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATGCTGACGCGTGAACAATGCTTCTGGG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATGCTGACGCGTGAACAATGCTTCTGGG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGATGCTGTCAGGTCGTAAATCACTGCATATTCG 1320

Db 1261 TCAGGAGCCATCGGAAGCTGTGATGCTGTGCAGGTCGTAAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATACGGTT 1380
Qy 1381 CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGCTGTATAATGTGGA 1440
Db 1381 CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGCTGTATAATGTGGA 1440
Qy 1441 ATTCTGAGCGGATAACAATTTTCACAGGAAACAGCCAGTCCGTTTATAGGTTTTCACGA 1500
Db 1441 ATTCTGAGCGGATAACAATTTTCACAGGAAACAGCCAGTCCGTTTATAGGTTTTCACGA 1500
Qy 1501 GCATTTCAACCAACAGGACCATAGATTATGAAATCTGAAGAGTAACTGGTAAATCTGG 1560
Db 1501 GCATTTCAACCAACAGGACCATAGATTATGAAATCTGAAGAGTAACTGGTAAATCTGG 1560
Qy 1561 ATTAAACGGGATAAAGCTATACCGTCTCGCTGAGTTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATAAAGCTATACCGTCTCGCTGAGTTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATAAACTGGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATAAACTGGAAGAGAAATTCGCCACAGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGCAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTACGTTACACCGCAAGCTGATTCCTTACCAGTCTGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTACACCGCAAGCTGATTCCTTACCAGTCTGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCAAACCCGCCAAAACCTGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCAAACCCGCCAAAACCTGGAA 1920
Qy 1921 GAGATCCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACGGCAAGTACGACATTAAGAGCGTGGCGTGGATAACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACGGCAAGTACGACATTAAGAGCGTGGCGTGGATAACGCTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACATGATGACACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACATGATGACACACCGGATTAC 2160
Qy 2161 TCCATCGCAAGAGCTGCCCTTTAATAAGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Db 2161 TCCATCGCAAGAGCTGCCCTTTAATAAGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACCGAGGAAAGTAAATGTTGTTAAACCGTACTGCCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCGAGGAAAGTAAATGTTGTTAAACCGTACTGCCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTTCTGCTGCTGAGCGCAGGATTAATAACCGCGCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTCTGCTGCTGAGCGCAGGATTAATAACCGCGCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGTTCCTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGTTCCTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400

841 TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db |||||
841 TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy |||||
901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCACTCAAAACAGGATTTTC 960
Db |||||
901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCACTCAAAACAGGATTTTC 960
Qy |||||
961 GCCTGCTGGGCAACACGAGCTGGACCGCTTCTGCGCACTCTCTCAGGCCAGCGGTGA 1020
Db |||||
961 GCCTGCTGGGCAACACGAGCTGGACCGCTTCTGCGCACTCTCTCAGGCCAGCGGTGA 1020
Qy |||||
1021 AGGCAATCAGCTGTTGGCCGCTCTCACTGTGTAAGAAAGAAACACACCTGCGGCCAATA 1080
Db |||||
1021 AGGCAATCAGCTGTTGGCCGCTCTCACTGTGTAAGAAAGAAACACACCTGCGGCCAATA 1080
Qy |||||
1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTT 1140
Db |||||
1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTT 1140
Qy |||||
1141 CCCGACTGGGAAGCGGCGAGTGAGCGAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Db |||||
1141 CCCGACTGGGAAGCGGCGAGTGAGCGAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Qy |||||
1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCAACCAATGCTTCTGGCG 1260
Db |||||
1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCAACCAATGCTTCTGGCG 1260
Qy |||||
1261 TCAGGACGCCATCGGAAGCTGTGATGCTGTGAGGTGCTGAATCACTGCATAATTCG 1320
Db |||||
1261 TCAGGACGCCATCGGAAGCTGTGATGCTGTGAGGTGCTGAATCACTGCATAATTCG 1320
Qy |||||
1321 TGTGCTCAAGCGCACCTCCCGTTCTGATAATGTTTTGCGCGGACATCATACCGGTT 1380
Db |||||
1321 TGTGCTCAAGCGCACCTCCCGTTCTGATAATGTTTTGCGCGGACATCATACCGGTT 1380
Qy |||||
1381 CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db |||||
1441 ATTGTGAGCGGATAACAAATTTACACAGGAACAGCCAGTCCGTTTAGGTGTTTTTCAG 1500
Qy |||||
1441 ATTGTGAGCGGATAACAAATTTACACAGGAACAGCCAGTCCGTTTAGGTGTTTTTCAG 1500
Db |||||
1501 GCACCTTCAACAAAGGACCATAGATTATGAAAATCTGAAGAGTAAACTGGTAATCTGG 1560
Qy |||||
1501 GCACCTTCAACAAAGGACCATAGATTATGAAAATCTGAAGAGTAAACTGGTAATCTGG 1560
Db |||||
1561 ATTAACCGCGGATAAAGCTATTAACGGTCTCGCTGAAGTTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy |||||
1561 ATTAACCGCGGATAAAGCTATTAACGGTCTCGCTGAAGTTCGGTAAAGAAATTCGAGAAAGAT 1620
Db |||||
1621 ACCGGAAATTAAGTCAACGGTTGAGCATCCGATAACTGGAAGAGAAATTCACAGGTT 1680
Qy |||||
1621 ACCGGAAATTAAGTCAACGGTTGAGCATCCGATAACTGGAAGAGAAATTCACAGGTT 1680
Db |||||
1681 GCGCAACTGCGCATGGCCCTGACATTTATCTTGGGCAACACGACCGCTTGGTGGCTAC 1740
Qy |||||
1681 GCGCAACTGCGCATGGCCCTGACATTTATCTTGGGCAACACGACCGCTTGGTGGCTAC 1740
Db |||||
1741 GCTCAATCTGCGCTGTTGGTGAATCAACCGCGAACAAGCGTTTCCAGGACAAAGCTGTAT 1800
Qy |||||
1741 GCTCAATCTGCGCTGTTGGTGAATCAACCGCGAACAAGCGTTTCCAGGACAAAGCTGTAT 1800
Db |||||
1801 CCGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATGCTTACCGATCGCTGTT 1860
Qy |||||
1801 CCGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db |||||
1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCGCGCAAAACCTGGGAA 1920
Qy |||||
1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCGCGCAAAACCTGGGAA 1920
Db |||||

1921 GAGATCCGCGCTGGATATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy |||||
1921 GAGATCCGCGCTGGATATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db |||||
1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTCTGACCGGGGTTATCGGTTCAAG 2040
Qy |||||
1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTCTGACCGGGGTTATCGGTTCAAG 2040
Db |||||
2041 TATGAAACCGGCAAGTACGACATTAAGAGCGTGGCGTGGATAAACGCTGCGCGAAAGCG 2100
Qy |||||
2041 TATGAAACCGGCAAGTACGACATTAAGAGCGTGGCGTGGATAAACGCTGCGCGAAAGCG 2100
Db |||||
2101 GGTCTGACCTTCTCGTTGACCTGATTTAAACAAACACATGATGACGACACCGATTAC 2160
Qy |||||
2101 GGTCTGACCTTCTCGTTGACCTGATTTAAACAAACACATGATGACGACACCGATTAC 2160
Db |||||
2161 TCCATCGCAGAAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Qy |||||
2161 TCCATCGCAGAAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db |||||
2221 GCATGCTCAACATCGACACAGCAAGTAATATGTTGTAACGGTACTGCGCACTTC 2280
Qy |||||
2221 GCATGCTCAACATCGACACAGCAAGTAATATGTTGTAACGGTACTGCGCACTTC 2280
Db |||||
2281 AAGGTCACCACTCCAAACCGTTTGGCGTCTGAGCGCAGGTATTTAACGCGCCAGT 2340
Qy |||||
2281 AAGGTCACCACTCCAAACCGTTTGGCGTCTGAGCGCAGGTATTTAACGCGCCAGT 2340
Db |||||
2341 CCGAAACAAAGAGCTGCGCAAAAGATTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Qy |||||
2341 CCGAAACAAAGAGCTGCGCAAAAGATTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db |||||
2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy |||||
2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db |||||
2461 TTGCGCAAAAGATCCACGTATTGCGCGCACCATCGAAACCGCCAGAAAGGTGAATCATG 2520
Qy |||||
2461 TTGCGCAAAAGATCCACGTATTGCGCGCACCATCGAAACCGCCAGAAAGGTGAATCATG 2520
Db |||||
2521 CCGAAACATCCGAGATGTCGGCTTTCTGATATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Qy |||||
2521 CCGAAACATCCGAGATGTCGGCTTTCTGATATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db |||||
2581 GCCAGCGCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
Qy |||||
2581 GCCAGCGCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
Db |||||
2641 AACAAACAAACAAATTAACAACTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Qy |||||
2641 AACAAACAAACAAATTAACAACTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db |||||
2701 GGATCTCTTCTCTGTCGCGCCAGGGCGCTCGAGCCGCGGGAGAAAGCCCTATGCTTGT 2760
Qy |||||
2701 GGATCTCTTCTCTGTCGCGCCAGGGCGCTCGAGCCGCGGGAGAAAGCCCTATGCTTGT 2760
Db |||||
2761 CCGGAATGTTGTAAGTCTTCTCTCAGAGCTCTCACTGTCGCGCCACAGCGTACCAC 2820
Qy |||||
2761 CCGGAATGTTGTAAGTCTTCTCTCAGAGCTCTCACTGTCGCGCCACAGCGTACCAC 2820
Db |||||
2821 ACGGGTGAACAAACCGTATAAATGCGCAGAGTGGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
Qy |||||
2821 ACGGGTGAACAAACCGTATAAATGCGCAGAGTGGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
Db |||||
2881 CTGGTGGCCATCAACCGACTCATCTGCGGAGAACCCATAAATGTCAGAAATGTGGC 2940
Qy |||||
2881 CTGGTGGCCATCAACCGACTCATCTGCGGAGAACCCATAAATGTCAGAAATGTGGC 2940
Db |||||
2941 AAGTCTTTCTCTCGGTCTGACAAATCTGTCGCGGACCAACGTAATCTCACACCGGGAGAG 3000
Qy |||||
2941 AAGTCTTTCTCTCGGTCTGACAAATCTGTCGCGGACCAACGTAATCTCACACCGGGAGAG 3000
Db |||||
3001 CCCTATGCTGTCCGGAATGTGGTAAGTCTTTCAGCCGAGCGATAACTGTGTGCGCCAC 3060
Qy |||||

|||||
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCCCTTCAGCCGAGCGATAAACCTGGTGGCCAC 3060
Qy 3061 CAGCGTACCCACACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Qy 3121 CAGCGCGGCACCTGGCCAGCCCATCAACGCACTCATCTACTGGGAGAACCATACAAATGT 3180
Db 3121 CAGCGCGGCACCTGGCCAGCCCATCAACGCACTCATCTACTGGGAGAACCATACAAATGT 3180
Qy 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGSCACCAAGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGSCACCAAGTACTCAC 3240
Qy 3241 ACCCGTAAAAAATAGTGGCCAGCGCCGACAGTACCCGCTACGAGTTCCGGACTACGCT 3300
Db 3241 ACCCGTAAAAAATAGTGGCCAGCGCCGACAGTACCCGCTACGAGTTCCGGACTACGCT 3300

RESULT 3
AX202430 3300 bp DNA linear PAT 30-AUG-2001
LOCUS Sequence 18 from Patent WO0152620.
DEFINITION AX202430
ACCESSION AX202430
VERSION AX202430.1 GI:15392178
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 18 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
Location/Qualifiers
source
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-Ap3 and zinc finger protein
ZFAPp3"

ORIGIN
Query Match 97.8%; Score 3226.4; DB 6; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCGGAAGAGA 60
Qy 61 GTCAAATTCAGGTGTGAATGTGAACCGTAACGTTATACGATGTCCGAGATATGCG 120
Db 61 GTCAAATTCAGGTGTGAATGTGAACCGTAACGTTATACGATGTCCGAGATATGCG 120
Qy 121 GTGTCCTTATCAGACCGTTTCCCGGTGTGAACCGAGCCAGCCAGTTCTCGGAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGGTGTGTGAACCGAGCCAGCCAGTTCTCGGAAAA 180
Qy 181 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCGCCAACCCGCTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCGCCAACCCGCTGGCAC 240
Qy 241 AACAACTGCGCGGCAAAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTGTGGCCCTGC 300
Db 241 AACAACTGCGCGGCAAAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTGTGGCCCTGC 300
Qy 301 ACGCCCGTTCGCAAAATGTTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGG 360
Db 301 ACGCCCGTTCGCAAAATGTTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGG 360

Qy 361 TGGTGGTGTGATCGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACATC 420
Db 361 TGGTGGTGTGATCGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACATC 420
Qy 421 TTCTCGGCAACCGGTCAAGTGGGCTGATCAATACTATCCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGGCAACCGGTCAAGTGGGCTGATCAATACTATCCGCTGGATGACAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTCGACAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTCGACAGA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGGGCTGGCGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGGGCTGGCGCTGGAGCATC 600
Qy 601 TGGTGCATTCGGGTCAACAGCAAAATCGGCTGTTAGCGGGGCCATTAAAGTTCTGTCGG 660
Db 601 TGGTGCATTCGGGTCAACAGCAAAATCGGCTGTTAGCGGGGCCATTAAAGTTCTGTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAAACGGGAAAGCGACATGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAAACGGGAAAGCGACATGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGCGCATGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGCGCTGGCGCAA 840
Db 781 ATGAGGCGCATGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGCGCTGGCGCAA 840
Qy 841 TGGCGCGCATTACCGAGTCCGGGCTGGCGGTGGTGGCGATATCTCGTAGTGGGATAG 900
Db 841 TGGCGCGCATTACCGAGTCCGGGCTGGCGGTGGTGGCGATATCTCGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTGTAACACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTGTAACACCATCAACAGGATTTTC 960
Qy 961 GCCTGTGGGGCAACACAGCGTGGACCGCTTGTGTCGAATCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGTGGGGCAACACAGCGTGGACCGCTTGTGTCGAATCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCGAGCTGGCACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCGAGCTGGCACGACAGGTTT 1140
Qy 1141 CCGACTGGAAAGCGGCGAGTGAGCGAACCAATTAATGTGAGTTAGTCTACTCATTTAG 1200
Db 1141 CCGACTGGAAAGCGGCGAGTGAGCGAACCAATTAATGTGAGTTAGTCTACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAACGCGTGCACCAATGTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAACGCGTGCACCAATGTTCTGGCG 1260
Qy 1261 TCAGCAGCCATTCGGAAGCTGTGGTATGGCTGTGAGTGTGAGTGTGTAATTCATCTGATTAATTCG 1320
Db 1261 TCAGCAGCCATTCGGAAGCTGTGGTATGGCTGTGAGTGTGAGTGTGTAATTCATCTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCAGCTCCCGTCTGGATTAATGTTTTGGCGGAGCATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCAGCTCCCGTCTGGATTAATGTTTTGGCGGAGCATCATACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTTCACAATTAATCATCGGCTCGTATATGTTGTTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTCACAATTAATCATCGGCTCGTATATGTTGTTGA 1440
Qy 1441 ATTGTGAGCGGATACAAATTTTCCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500

ORIGIN		2ZFPm4"	
Query Match		97.4%; Score 3215.2; DB 6; Length 3300;	
Best Local Similarity		98.4%; Pred. No. 0;	
Matches 3247; Conservative		0; Mismatches 53; Indels 0; Gaps 0;	
Qy	1	CGGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATAGCATATGCGCCCGGAAGAGA	60
Db	1	CGGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATAGCATATGCGCCCGGAAGAGA	60
Qy	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120
Db	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120
Qy	121	GTGCTCTTATCAGACCCGTTTCCCGCGTGTGAACACGCGCACGCTTTCTCGGAANA	180
Db	121	GTGCTCTTATCAGACCCGTTTCCCGCGTGTGAACACGCGCACGCTTTCTCGGAANA	180
Qy	181	CGCGGAAAAGTGAAGCGCGATGGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCNC	240
Db	181	CGCGGAAAAGTGAAGCGCGATGGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCNC	240
Qy	241	AACAACTGCGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGSCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGCGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGSCACCTCCAGTCTGGCCCTGC	300
Qy	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
Qy	361	TGGTGGTGTGATGTAGAACGAAAGCGGCTGAAAGCTGTAAAGCGGCGTGCACAATC	420
Db	361	TGGTGGTGTGATGTAGAACGAAAGCGGCTGAAAGCTGTAAAGCGGCGTGCACAATC	420
Qy	421	TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
Db	421	TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
Qy	481	TTGCTGTGGAACTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAGA	540
Db	481	TTGCTGTGGAACTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAGA	540
Qy	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
Qy	601	TGGTGGTGGTGCACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTCTGCTCGG	660
Db	601	TGGTGGTGGTGCACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTCTGCTCGG	660
Qy	661	CGCGTCTGGCTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGGCTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAAACGGGAAGCGACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAAACGGGAAGCGACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA	780
Qy	781	ATGAGGGCATCGTCCCACTGGATGCTGGTTCGCAACGATCAGATGGCGCTGGGGCAA	840
Db	781	ATGAGGGCATCGTCCCACTGGATGCTGGTTCGCAACGATCAGATGGCGCTGGGGCAA	840
Qy	841	TGCGGCCATTACCGAGTCCGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATAGC	900
Db	841	TGCGGCCATTACCGAGTCCGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATAGC	900
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAATCAACAGGATTTTC	960
Qy	961	GCCTGCTGGGGCAACACGCTGGACCGCTTGTGCAACTCTCTCAGGCCAGGCGGTGA	1020
Db	961	GCCTGCTGGGGCAACACGCTGGACCGCTTGTGCAACTCTCTCAGGCCAGGCGGTGA	1020
Qy	1021	AGGCAATCAGCTGTGTCCTCTCACTGCTGCAACTCTCTCAGGCCAGGCGGTGA	1080
Db	1021	AGGCAATCAGCTGTGTCCTCTCACTGCTGCAACTCTCTCAGGCCAGGCGGTGA	1080
Qy	1081	CGCAAAACCGCTCTCCCGCGGCTGGCCGATTCATTAATGCAGCTGGCAGCAGAGGTTT	1140
Db	1081	CGCAAAACCGCTCTCCCGCGGCTGGCCGATTCATTAATGCAGCTGGCAGCAGAGGTTT	1140
Qy	1141	CCGACTGGAAGGGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGTCACTCATTTAG	1200
Db	1141	CCGACTGGAAGGGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGTCACTCATTTAG	1200
Qy	1201	GCAAAATTTCTCATGTTTGACAGCTTATCATCGACTGCAACCGGTGCAACCAATGCTTCGGCG	1260
Db	1201	GCAAAATTTCTCATGTTTGACAGCTTATCATCGACTGCAACCGGTGCAACCAATGCTTCGGCG	1260
Qy	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATTAATTCG	1320
Db	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATTAATTCG	1320
Qy	1321	TGTGCTCAAGCGCGCACTCCCGTTCTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380
Db	1321	TGTGCTCAAGCGCGCACTCCCGTTCTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATATGTTGTGGA	1440
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATATGTTGTGGA	1440
Qy	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACGCGAGTCCGTTTAGGTGTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACGCGAGTCCGTTTAGGTGTTTTCACGA	1500
Qy	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAGTGAAGAGGTAATCTGGTAATCTGG	1560
Db	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAGTGAAGAGGTAATCTGGTAATCTGG	1560
Qy	1561	ATTAAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
Db	1621	ACCGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
Qy	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
Db	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
Db	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
Qy	1801	CGGTTTACCTGGGATGCGTACGTTTACACCGCAAGCTGATTGCTTACCGGATCGCTGTT	1860
Db	1801	CGGTTTACCTGGGATGCGTACGTTTACACCGCAAGCTGATTGCTTACCGGATCGCTGTT	1860
Qy	1861	GAAAGGTTATCGCTGATTATAACAAAGATCTGTCGCGAACCCCGCAAAACCTGGGAA	1920
Db	1861	GAAAGGTTATCGCTGATTATAACAAAGATCTGTCGCGAACCCCGCAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAAACCGCAAGTACGACATTAAGACGTTGGGCGTGGATTAACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAAACCGCAAGTACGACATTAAGACGTTGGGCGTGGATTAACGCTGGCGCGAAAGCG	2100

QY 541 CACCCATCAAAGTATTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGCAATGGGTCACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTCGCAATGGGTCACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTCACAAACCAATGCAATGCTGA 780
DB 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTCACAAACCAATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCACTGGCATGTCTGGTTGCCAACGATCAGATGCGCTGGCGGCA 840
DB 781 ATGAGGGCATCGTTCCTCACTGGCATGTCTGGTTGCCAACGATCAGATGCGCTGGCGGCA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCAGATATCTCGGTAGTGGGATAG 900
DB 841 TCGCGGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCAGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGAGTGGAGCCGCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTCTGGGGCAACACGAGTGGAGCCGCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 1080
DB 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 1080
QY 1081 CGCAAAACGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
DB 1081 CGCAAAACGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTACTCACTCATTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTACTCACTCATTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAGCTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATTCGGAAGCTGTGGTATGGCTGTGAGCTCGTAATCACTGCATTAATTCG 1320
DB 1261 TCAGGAGCCATTCGGAAGCTGTGGTATGGCTGTGAGCTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTTCCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGGTT 1380
DB 1321 TGTTCCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAATTAATCATCGGCTCGTATATGTGTGA 1440
DB 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAATTAATCATCGGCTCGTATATGTGTGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTCACAGAAACAGCCAGTCCGTTTAGTGTTTTTCAGCA 1500
DB 1441 ATTGTAGCGGATAACAATTTTCACAGAAACAGCCAGTCCGTTTAGTGTTTTTCAGCA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAACCTGAAGAGGTAAATCGGTAATCTCG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAACCTGAAGAGGTAAATCGGTAATCTCG 1560
QY 1561 ATTTAAGCGGATAAAGGCTTAAACCGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTTAAGCGGATAAAGGCTTAAACCGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCAACCGTTGAGATCCGGATAAATCTGGAAGAGAAATTTCCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGATCCGGATAAATCTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB 1681 CGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGACAAGAGGTTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGACAAGAGGTTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAACGAAAGGTAAGAGCGCGCTGATTTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATAAAGAACTGAAACGAAAGGTAAGAGCGCGCTGATTTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTACCTGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
DB 2041 TATGAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAAACAACATGAATGACAGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAAACAACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTAATAAGGCGAAGACAGCGATGACCATCAACGGCCCGTGG 2220
DB 2161 TCCATCGCAGAGCTGCCCTTTAATAAGGCGAAGACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACGCAAGTGAATTAATGTGTAAACGGTACTGCGGACCTTC 2280
DB 2221 GCATGCTCAACATCGACACGCAAGTGAATTAATGTGTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACACCATCAAAACCGTTTGGTGGCGTCTGAGCGCAGGTATTAACGCCGCCAGT 2340
DB 2281 AAGGTCACACCATCAAAACCGTTTGGTGGCGTCTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CGAACAAGAGCTGGGAAAAGAGTTTCTCGAAAACCTATCTGCTGACCTGATGAAGTCTG 2400
DB 2341 CGAACAAGAGCTGGGAAAAGAGTTTCTCGAAAACCTATCTGCTGACCTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
DB 2401 GAAGCGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
QY 2461 TTGCGAAAGATTCACGATTTTGGCGCCACCATGGAAGAAACGCGGAGAGGTGAATCATG 2520
DB 2461 TTGCGAAAGATTCACGATTTTGGCGCCACCATGGAAGAAACGCGGAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
DB 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGACACTTAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGACACTTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
DB 2641 AACAAACAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGTTGGCCCAAGCGGCCCTTCGAGAGCCCGGGGAGAACCCCTATGCTGT 2760

```
Db 2701 GGATCCTCTCTCTGTCGCCAGCGCGCTCGAGCCGCGGAGAGCCCTATCTGT 2760
Qy 2761 CCGGAATGGAAGTCTCTCTCAGAGCTCTCAGCTGTCGCGCCAGCGGTACCCAC 2820
Db 2761 CCGGAATGGAAGTCTCTCTCAGAGCTCTCAGCTGTCGCGCCAGCGGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGTCGCGAGTGGCGCAAAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGTCGCGAGTGGCGCAAAATCTTTAGCCAGCGGTCTCC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Qy 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCGCGCACCAACGCTACTCACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCGGTCTGACAACTCTGTCGCGCACCAACGCTACTCACCGGGGAGAG 3000
Qy 3001 CCCTATGTTGTCGGAATGTTGTAAGTCTTTCAGCCGAGCGATAAATCTGTCGCCAC 3060
Db 3001 CCCTATGTTGTCGGAATGTTGTAAGTCTTTCAGCCGAGCGAGTCTTCTGTCGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGTCGCGAGTGGCGCAAAATCTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGTCGCGAGTGGCGCAAAATCTTTAGT 3120
Qy 3121 CAGGCGCGCCACCTGCGGAGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 GACTGCGCGACCTTGTCTGCGCCTCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Qy 3181 CCAGATGTGCAAGTCTTTCTCGGTGTCGCAATCTGTCGCGCACCAACGTAAGTCTCAC 3240
Db 3181 CCAGATGTGCAAGTCTTTCTCGGTGTCGCAATCTGTCGCGCACCAACGTAAGTCTCAC 3240
Qy 3241 ACCGTAATAAATAGTGGCGCGCGCGCGAGTACCGGTACGAGCTTCGGAGTACGCT 3300
Db 3241 ACCGTAATAAATAGTGGCGCGCGCGCGAGTACCGGTACGAGCTTCGGAGTACGCT 3300

RESULT 6
AX284144 LOCUS 7373 bp DNA linear PAT 20-NOV-2001
DEFINITION Sequence 1 from Patent WO0178756.
ACCESSION AX284144
VERSION AX284144.1 GI:17044832
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1.
AUTHORS Wiedersanders, B. and Maubach, G.
TITLE Agent for postoperative use after the removal of bone tumours
JOURNAL Patent: WO 0178756-A 1 25-OCT-2001;
Depuy Biotech Jena GmbH (DE)
FEATURES
source
1. .7373
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Expressionsvektor pMalc2cchbnp2"
1528..3453
/note="unnamed protein product; /produkt=Fusionsprotein
MBP Cystatin C BMP-2"
/codon_start=1
/transl_table=11
/protein_id="CAD12160.1"
/db_xref="GI:17044833"
/translation="MKIEEKLIWINGDKYNGLAEVKKFKDGIKIVTVEHPDKL
EEKFPQVATGDDPDIIFWAHDFRFGTAQGLLAEITPDKAFODKILYPTFVADRVYNG
KLIAADYAFBALSLIYNKDLPPPKTWEEIPALDKELKAKSKALMFLNLOEYFPTWP
LIAADGGVAFKNGYKIDVGVNAGAKAGLTFVLDTIKNKHMNADTDSIAEAAF
NKGETAMTINGPWAWSNIDTISKVNYGVTVLPTFKGQPSKFPFVGLSAGINAASPNKEL
```

```
misc_feature
/note="Faktor Xa Protease /position=388 (AMINO
ACIDsequenz)"
3067..3108
/misc_feature
/note="/spacer zwischen Cystatin C und BMP-2"
ORIGIN
Query Match 82.0%; Score 2705.8; DB 6; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGTCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGTCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
Qy 61 GTCAAATTCAGGGTGGTGAATGTGAAACACAGTAACTTATACGATGTCCGAGAGTATCCG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAAACACAGTAACTTATACGATGTCCGAGAGTATCCG 120
Qy 121 GTGCTCTTTATCAGACCGTTTCCGCGGTGTGAACCAAGCCAGCCAGCGTTCCTGCGAAAA 180
Db 121 GTGCTCTTTATCAGACCGTTTCCGCGGTGTGAACCAAGCCAGCCAGCGTTCCTGCGAAAA 180
Qy 181 CCGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCACCGCGGTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCACCGCGGTGGCAC 240
Qy 241 AACAACTGGCGGCAACAGTCTGATGCGCGTTCGACCTCCAGTCTGCGCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCTGATGCGCGTTCGACCTCCAGTCTGCGCCTGC 300
Qy 301 ACCGCGCGTCGCAAAATGTCGCGCGCATTAATTAATTCGCGCGCATCAACTGGGTGCCAGC 360
Db 301 ACCGCGCGTCGCAAAATGTCGCGCGCATTAATTAATTCGCGCGCATCAACTGGGTGCCAGC 360
Qy 361 TGGTGTGTGTCGATGTCGAGAACGAGCGCGTCTGAAAGCTGTAAGCGCGGTGCACAATC 420
Db 361 TGGTGTGTGTCGATGTCGAGAACGAGCGCGTCTGAAAGCTGTAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTAATTCGCGTGGATGACAGAGTGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTAATTCGCGTGGATGACAGAGTGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGACATAATTTTCGCGCGTATTTCTTGTATGTCCTGAGCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGACATAATTTTCGCGCGTATTTCTTGTATGTCCTGAGCAGA 540
Qy 541 CACCCATCAACAGTATTAATTTTCCTCCATGAAGAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTTCCTCCATGAAGAGCGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGTGTCGATGTCACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCCTGTCGG 660
Db 601 TGGTGTGTCGATGTCACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCCTGTCGG 660
Qy 661 CCGGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CCGAACCGGAAGCGGAGTGGCATGTCGCGTTCCTCAACAAACCAATGCTGA 780
Db 721 CCGAACCGGAAGCGGAGTGGCATGTCGCGTTCCTCAACAAACCAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCTCCACTGCGATGTCGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCTCCACTGCGATGTCGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
```

QY	841	TGCGGCGCATTAACCGAGTCCGGGCTGCGCGTTGGTGGCGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGGCGCATTAACCGAGTCCGGGCTGCGCGTTGGTGGCGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACGAGATTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACGAGATTTC	960
QY	961	GCCTCTGGGCAAAACACGCGTGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020
Db	961	GCCTCTGGGCAAAACACGCGTGAGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTGTCGCCGTTCTCACTGGTGAAAGAAAACCAACCCCTGGCGGCCAATA	1080
Db	1021	AGGGCAATCAGCTGTGTCGCCGTTCTCACTGGTGAAAGAAAACCAACCCCTGGCGGCCAATA	1080
QY	1081	CGCAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAAATGCAAGTGGCAGACAGGTTT	1140
Db	1081	CGCAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAAATGCAAGTGGCAGACAGGTTT	1140
QY	1141	CCCGACTGGAAGCGGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG	1200
Db	1141	CCCGACTGGAAGCGGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG	1200
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATGCACTGCAAGTGGCAGCAATGCTTCTGGCG	1260
Db	1201	GCACAATTTCTCATGTTTGACAGCTTATCATGCACTGCAAGTGGCAGCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCGATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTCG	1320
Db	1261	TCAGGACGCGATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT	1380
QY	1381	CTGGCAATATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440
Db	1381	CTGGCAATATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440
QY	1441	ATTGTAGCGGATACAAATTTACACAGGAAACAGCGAGTCGGTTAGTGTTTTCACGA	1500
Db	1441	ATTGTAGCGGATACAAATTTACACAGGAAACAGCGAGTCGGTTAGTGTTTTCACGA	1500
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAACTGGAAGAGTTAAACTGGTAATCTGG	1560
Db	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAACTGGAAGAGTTAAACTGGTAATCTGG	1560
QY	1561	ATTAAACGCGATAAAGGCTATAACGGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGCGATAAAGGCTATAACGGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680
Db	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680
QY	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACAGCGCTTTGGTGGCTAC	1740
Db	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACAGCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGGCTGTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAAGTGTAT	1800
Db	1741	GCTCAATCTGGGCTGTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAAGTGTAT	1800
QY	1801	CGTTTACCTGGATGCGGTAAGTCAACCGGAAAGCTGATTGCTTACCCGATCGCTGTT	1860
Db	1801	CGTTTACCTGGATGCGGTAAGTCAACCGGAAAGCTGATTGCTTACCCGATCGCTGTT	1860
QY	1861	GAAGGTTATCCTGATTTATACAAAGATCTGCTCGCAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAAGGTTATCCTGATTTATACAAAGATCTGCTCGCAACCCGCCAAAACCTGGGAA	1920

QY	1921	GAGATCCGCGCGCTGGATAAAGAACTGAAGCGAAAGAGTAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGCGCGCTGGATAAAGAACTGAAGCGAAAGAGTAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAAACGTAATTTACCTGCGCGCTGATGCTGCTGACGGGGGTTATGCGGTTCAAG	2040
Db	1981	CTGCAAGAAACGTAATTTACCTGCGCGCTGATGCTGCTGACGGGGGTTATGCGGTTCAAG	2040
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGGTGGATAACGCTGGCGGAAACGG	2100
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGGTGGATAACGCTGGCGGAAACGG	2100
QY	2101	GCTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATCAATGACAGACCGATTAC	2160
Db	2101	GCTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATCAATGACAGACCGATTAC	2160
QY	2161	TCCATCGCAGAAAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAAGCGCCGCTGG	2220
Db	2161	TCCATCGCAGAAAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAAGCGCCGCTGG	2220
QY	2221	GCATGGTCCAAACATCGACACCAAGTGAATTAATGGTGTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGGTCCAAACATCGACACCAAGTGAATTAATGGTGTAAACGGTACTGCCGACCTTC	2280
QY	2281	AAGGTCAACCATCAAAACCGTTGCTGGCGTGTGAGCGAGGTATTAACCGCCCGAGT	2340
Db	2281	AAGGTCAACCATCAAAACCGTTGCTGGCGTGTGAGCGAGGTATTAACCGCCCGAGT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACCTGATGAAGGCTCG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACCTGATGAAGGCTCG	2400
QY	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTCCGTAGCGCTGAAGTCTTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTCCGTAGCGCTGAAGTCTTTACGAGGAAGAG	2460
QY	2461	TTGGGAAAGATTCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATTCATG	2520
Db	2461	TTGGGAAAGATTCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATTCATG	2520
QY	2521	CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCTGCTGCTGCTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCTGCTGCTGCTGATCAACGCC	2580
QY	2581	GCCAGCGTCTGACGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGTCTGACGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
QY	2641	AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC	2700
Db	2641	AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC	2700
QY	2701	GGATCCTCTCTCTGTCGGCCGAGCGCGCTTCGAGGCCCGG	2741
Db	2701	GGATCCTCTCTCTGTCGGCCGAGCGCGCTTCGAGGCCCGG	2741

RESULT 7

AX172306

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AX172306 8101 bp DNA linear PAT 03-JUL-2001
Sequence 5 from Patent WO0144467.

AX172306
AX172306.1 GI:14597487

synthetic construct
synthetic construct
other sequences; artificial sequences.

1
Goubin-Gramatica,F., Ducommun,B. and Prevost,G.
Method for obtaining human cdc25 phosphatases and method for
identifying human cdc25 phosphatase modulators
Patent: WO 0144467-A 5 21-JUN-2001;
SOCIETE DE CONSEILS DE RECHERCHES ET D'APPLICATIONS SCIENTIFIQUES

FEATURES		Location/Qualifiers	ORIGIN
source			
(FR)			
1. .8101			
/organism="synthetic construct"			
/mol_type="unassigned DNA"			
/db_xref="taxon:32630"			
/note="souche JM109 / pMAL - Hs Cdc25C"			
Query Match		82.0%; Score 2705.8; DB 6; Length 8101;	
Best Local Similarity		99.9%; Pred. No. 0;	
Matches 2707; Conservative		0; Mismatches 2; Indels 0; Gaps 0;	
Qy	1	CGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGAGCCCGGAAGAGA	60
Db	1	CGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGAGCCCGGAAGAGA	60
Qy	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCCAGAGTATGCCG	120
Db	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCCAGAGTATGCCG	120
Qy	121	GTGTCCTTTATCAGACCGTTTCCCGGTGGTGAACCGCCAGCCAGCTTTCTCGGAAA	180
Db	121	GTGTCCTTTATCAGACCGTTTCCCGGTGGTGAACCGCCAGCCAGCTTTCTCGGAAA	180
Qy	181	CGGGGAAAAAGTGAACCGCGATGGCGAGCTGAATTCATTTCCCAACCGCGTGGCAC	240
Db	181	CGGGGAAAAAGTGAACCGCGATGGCGAGCTGAATTCATTTCCCAACCGCGTGGCAC	240
Qy	241	AACAATCGCGGCAACAGTCGTGCTGATGGGGTGGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAATCGCGGCAACAGTCGTGCTGATGGGGTGGCCACCTCCAGTCTGGCCCTGC	300
Qy	301	ACGCGCGTGCATAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGGTGCCAGCG	360
Db	301	ACGCGCGTGCATAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGGTGCCAGCG	360
Qy	361	TGTTGGTGTGATGTAGAACGAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAATC	420
Db	361	TGTTGGTGTGATGTAGAACGAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAATC	420
Qy	421	TTCTCGGCAACCGGTCACTGAGTGAATTAATCTCGCGCGGATTAATCTCGCGCGGAT	480
Db	421	TTCTCGGCAACCGGTCACTGAGTGAATTAATCTCGCGCGGATTAATCTCGCGCGGAT	480
Qy	481	TTGCTGTGGAAGTGCCTGCATAATTTTCGGCGGTTATTTCTTGATGTCTCTGACCAGA	540
Db	481	TTGCTGTGGAAGTGCCTGCATAATTTTCGGCGGTTATTTCTTGATGTCTCTGACCAGA	540
Qy	541	CACCATCAACAGTATTATTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600
Db	541	CACCATCAACAGTATTATTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600
Qy	601	TGTTGCGAATTTGGGTCAACAGCAATTCGGTGTAGCGGGCCCAATTAAAGTTCTGCTCGG	660
Db	601	TGTTGCGAATTTGGGTCAACAGCAATTCGGTGTAGCGGGCCCAATTAAAGTTCTGCTCGG	660
Qy	661	CGGCTCTGCGTCTGGCTGGGTGCAATAATCTCACTCGCAATCAAAATTCACCCGATAG	720
Db	661	CGGCTCTGCGTCTGGCTGGGTGCAATAATCTCACTCGCAATCAAAATTCACCCGATAG	720
Qy	721	CGGAACGGGAAGGACGTGGAGTGGCATATCTCGGTGTTTCAACAAACATGCAATGCTGA	780
Db	721	CGGAACGGGAAGGACGTGGAGTGGCATATCTCGGTGTTTCAACAAACATGCAATGCTGA	780
Qy	781	ATGAGGCAATCGTTTCCCACTCGCATGTGTTGGCAACGATCAGATGGCGCTGGGCGCAA	840
Db	781	ATGAGGCAATCGTTTCCCACTCGCATGTGTTGGCAACGATCAGATGGCGCTGGGCGCAA	840
Qy	841	TGCGCGCATTAACGAGTCCGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT	900
Db	841	TGCGCGCATTAACGAGTCCGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT	900
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGATTTTC	960
Qy	961	GCTTGTGGGGCAAAACAGCGGTGGACCGCTTGTGTGCAAACTCTCTCAGGGGCGAGCGTGA	1020
Db	961	GCTTGTGGGGCAAAACAGCGGTGGACCGCTTGTGTGCAAACTCTCTCAGGGGCGAGCGTGA	1020
Qy	1021	AGGCAATCAGCTGTGTCCT	1080
Db	1021	AGGCAATCAGCTGTGTCCT	1080
Qy	1081	CGAAAACCGCT	1140
Db	1081	CGAAAACCGCT	1140
Qy	1141	CCGCAATCGAAAAGCGGCAAGTGAACGCAACGCAATTAATGTGAGTTAGCTCACTATTAG	1200
Db	1141	CCGCAATCGAAAAGCGGCAAGTGAACGCAACGCAATTAATGTGAGTTAGCTCACTATTAG	1200
Qy	1201	GCACAAATTCATGTTTTCAGAGCTTATCATCGAGTGCACGGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATTCATGTTTTCAGAGCTTATCATCGAGTGCACGGTGCACCAATGCTTCTGGCG	1260
Qy	1261	TCAGGCAAGCCATCGGAAGCTGTGATGGCTGTGCAAGTGCATAATCACTGCATAATTTCG	1320
Db	1261	TCAGGCAAGCCATCGGAAGCTGTGATGGCTGTGCAAGTGCATAATCACTGCATAATTTCG	1320
Qy	1321	TGTCGCTCAAGGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCCCGACATCAACCGGTT	1380
Db	1321	TGTCGCTCAAGGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCCCGACATCAACCGGTT	1380
Qy	1381	CTGCAAAATTTCTGAAATGAGCTGTGCAATTAATCATCGCTCGTAAATGTGTGGA	1440
Db	1381	CTGCAAAATTTCTGAAATGAGCTGTGCAATTAATCATCGCTCGTAAATGTGTGGA	1440
Qy	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCAGTCCGTTTTCAGTGTTCACGA	1500
Db	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCAGTCCGTTTTCAGTGTTCACGA	1500
Qy	1501	GCATTTCAACCAAGGACCATAGATTTGAAATCTGAAAGAGGTAAATCTGTTAAATCTGG	1560
Db	1501	GCATTTCAACCAAGGACCATAGATTTGAAATCTGAAAGAGGTAAATCTGTTAAATCTGG	1560
Qy	1561	ATTAAACGGGATTAAGCGCTTAAACCGTCTCGCTGAGTAAAGAAATTCGGAAGAT	1620
Db	1561	ATTAAACGGGATTAAGCGCTTAAACCGTCTCGCTGAGTAAAGAAATTCGGAAGAT	1620
Qy	1621	ACCGGAATTAAGTCAACCGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACACAGGTT	1680
Qy	1681	GGGCAACTGGGATGGCGCTGACATTTCTTGGGCAACAGCCGCTTGGTGGCTAC	1740
Db	1681	GGGCAACTGGGATGGCGCTGACATTTCTTGGGCAACAGCCGCTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCTGTGCTGAAATCACCCGGAACAAAGCGTTCAGGACCAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCTGTGCTGAAATCACCCGGAACAAAGCGTTCAGGACCAAGCTGTAT	1800
Qy	1801	CGTTTTACTCGGATGCCGTAACCGCAAGCTGATTTACCCGATCGCTGTT	1860
Db	1801	CGTTTTACTCGGATGCCGTAACCGCAAGCTGATTTACCCGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTACCTGGCGGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG	2040

[illegible]

/note="Plasmid pMAL-c2x"									
ORIGIN	Query Match	81.8%	Score 2701;	DB 6;	Length 6648;				
	Best Local Similarity	99.8%;	Pred. No. 0;						
	Matches 2704;	Conservative	0;	Mismatches	5;	Indels	0;	Gaps	0;
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGGCATGATAGGCCCGGGAAGAGA	60						
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGGCATGATAGGCCCGGGAAGAGA	60						
Qy	61	GTCAATTCAGGGTGGTGAATGTGAAACGAGTAAAGTTATACGATGTGCGAGAGTATGCGG	120						
Db	61	GTCAATTCAGGGTGGTGAATGTGAAACGAGTAAAGTTATACGATGTGCGAGAGTATGCGG	120						
Qy	121	GTGTCTCTTATCAGACACCGTTTCCCGCGTGGTGAACGAGGCCAGCACGCTTTCGCGAAAA	180						
Db	121	GTGTCTCTTATCAGACACCGTTTCCCGCGTGGTGAACGAGGCCAGCACGCTTTCGCGAAAA	180						
Qy	181	CGCGGGAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTACATTCCTCCAAACCGCTGGCAC	240						
Db	181	CGCGGGAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTACATTCCTCCAAACCGCTGGCAC	240						
Qy	241	AACAACTGGCGGGGAAAAAGTCGTTGCTGATTTGGCGTTGCCACTCTCCAGTCTGGGCCCTTGC	300						
Db	241	AACAACTGGCGGGGAAAAAGTCGTTGCTGATTTGGCGTTGCCACTCTCCAGTCTGGGCCCTTGC	300						
Qy	301	ACGCGCGCTGCGCAAAATGTGCGCGCGATTAAATCTCGCGCGCATCAACTGGGTGCCAGCG	360						
Db	301	ACGCGCGCTGCGCAAAATGTGCGCGCGATTAAATCTCGCGCGCATCAACTGGGTGCCAGCG	360						
Qy	361	TGTTGGTGTTCGATGTGAAACGAAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC	420						
Db	361	TGTTGGTGTTCGATGTGAAACGAAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC	420						
Qy	421	TTCTTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGATGATGACACGAGATGCCA	480						
Db	421	TTCTTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGATGATGACACGAGATGCCA	480						
Qy	481	TTGCTGTGGAGCTGCCCTGCACTAATGTTCCGGCGTTAATTTCTTCATGCTCTCTGACCCAGA	540						
Db	481	TTGCTGTGGAGCTGCCCTGCACTAATGTTCCGGCGTTAATTTCTTCATGCTCTCTGACCCAGA	540						
Qy	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600						
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600						
Qy	601	TGTTGCGCATTTGGGTCAACAGCAAAATCGCGCTGTATTAGCGGGGCCCATTAAGTTCTGTCTCGG	660						
Db	601	TGTTGCGCATTTGGGTCAACAGCAAAATCGCGCTGTATTAGCGGGGCCCATTAAGTTCTGTCTCGG	660						
Qy	661	CGCGTCTCGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGCGTCTCGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Qy	721	CGGAAACGGGAAGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTCA	780						
Db	721	CGGAAACGGGAAGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTCA	780						
Qy	781	ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGGCCAAACGATCAGATGCGCTGGGCGCA	840						
Db	781	ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGGCCAAACGATCAGATGCGCTGGGCGCA	840						
Qy	841	TGCGGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGCGGATATCTCGGTAGTGGGATACG	900						
Db	841	TGCGGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGCGGATATCTCGGTAGTGGGATACG	900						
Qy	901	ACGATACCGAAGACAGCTCATGTTATATATCCCGCGGTTAACCCACCATCAAAACAGGATTTTC	960						
Db	901	ACGATACCGAAGACAGCTCATGTTATATATCCCGCGGTTAACCCACCATCAAAACAGGATTTTC	960						
Qy	961	GCCTGCTGGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGGGTGA	102						

Db 961 GCCTGCTGGGCAACACGAGCTGGAGCCGCTTGCTGCAACTCTCTCAGGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCGCTCCTCAGCTGGTGAAGAAACACACCTTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCGCTCCTCAGCTGGTGAAGAAACACACCTTGGCGCCCAATA 1080
Qy 1081 GCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCAGAGGTTT 1140
Db 1081 GCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCAGAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTACTCACTCAATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTACTCACTCAATTAG 1200
Qy 1201 GCACAATCTCATGTTTGCACAGCTTATCATGACCTGCACGCTGCACCAATGCTTCTGCG 1260
Db 1201 GCACAATCTCATGTTTGCACAGCTTATCATGACCTGCACGCTGCACCAATGCTTCTGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGTGGTATGCTGTGAGCTGTAATCACTGCATAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGCTGTGAGCTGTAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTCTGGAATATGTTTTCGCGGAGCATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTCTGGAATATGTTTTCGCGGAGCATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTTTCACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTTTCACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1440
Qy 1441 ATTGTAGCGGATTAACAAATTTACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATTAACAAATTTACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCATTTCACCAACAGGACCATAGATTATGAACCTGAGAGGTAACCTGCTAATCTGG 1560
Db 1501 GCATTTCACCAACAGGACCATAGATTATGAACCTGAGAGGTAACCTGCTAATCTGG 1560
Qy 1561 ATTAACCGCGATAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAGGTT 1680
Qy 1681 GCGCAACTGCGGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGTGGTGAATCAACCCGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGTGGTGAATCAACCCGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGTACGTTTACACGGCAAGCTGATTGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTTACACGGCAAGCTGATTGCTTACCGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCGCAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCGCAACCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCGGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCTGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCTGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAAGACGCTGGCGTGGATTAACGCTGCGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAAGACGCTGGCGTGGATTAACGCTGCGCGGAAAGCG 2100

Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAACAAACACATGATGACAGACCCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAACAAACACATGATGACAGACCCGATTAC 2160
Qy 2161 TCCATCCGAGAAGCTGCTTTTAAATAGGCGGAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAAGCTGCTTTTAAATAGGCGGAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTTATGGTGTAAACGGTACTTGCACACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTTATGGTGTAAACGGTACTTGCACACCTTC 2280
Qy 2281 AAGGTTCAACATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGGCCCGT 2340
Db 2281 AAGGTTCAACATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGGCCCGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGAGGTTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGAGGTTCTG 2400
Qy 2401 GAAGCGGTTTAAAGCAAAACCCGCTGGTGGTGGCTAGCGTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTTAAAGCAAAACCCGCTGGTGGTGGCTAGCGTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAGATCCAGTATTCGGCCACCATCGAAACCGCCAGAAAGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCAGTATTCGGCCACCATCGAAACCGCCAGAAAGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGATATGCGTGGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGATATGCGTGGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCAGCGCTGCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCAGCGCTGCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATTAACAAATTAACAAACAAACCTCGGATCGAGGAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAAACAAATTAACAAATTAACAAACAAACCTCGGATCGAGGAGGATTTTCAAGATTC 2700
Qy 2701 GGATCTCT 2709
Db 2701 GGATCTCT 2709

RESULT 9
A91965
LOCUS
DEFINITION
A91965
Accession
A91965.1
GI:6740811
unidentified
unclassified.
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN

A91965
Sequence 1 from Patent WO9822496.
A91965.1
GI:6740811
unidentified
unclassified.
Attwood, M.R. and Hurst, D.N.
ANTIVIRAL PEPTIDE DERIVATIVES
Patent: WO 9822496-A 1 28-MAY-1998;
HOFFMANN LA ROCHE (CH)
Location/Qualifiers
1. .7475
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGACACCATCATGATGGTGCAGAAACCTTTCCGGGTATGCGGTATGATAGCGCCCGGAAGAGA 60

|||||
1 CCGACACCATCGAATGGTGTCAAAACCTTTCGCGGTATGSCATGATAGCGCCCGGAAGAGA 60
61 GTCAATTACAGGTTGGTGAATGTGAACACCAAGTAACTGATACGATGTGCGAGAGTATGCCG 120
61 GTCAATTACAGGTTGGTGAATGTGAACACCAAGTAACTGATACGATGTGCGAGAGTATGCCG 120
121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGCGCAGCCACGTTTCTGCGAANA 180
121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGCGCAGCCACGTTTCTGCGAANA 180
181 CCGCGGAAAAGTGAAGCGGATGGCGAGCTGAAATTTACATTTCCCAACCGCGTGGCAC 240
181 CCGCGGAAAAGTGAAGCGGCGATGGCGAGCTGAAATTTACATTTCCCAACCGCGTGGCAC 240
241 AACAACTCGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
301 AGCGCCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360
301 AGCGCCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360
361 TGGTGTGTTCGATGTAGAACCAAGCGGCTCGAAGCCTGTAAAGCGGCGTGCACAATC 420
361 TGGTGTGTTCGATGTAGAACCAAGCGGCTCGAAGCCTGTAAAGCGGCGTGCACAATC 420
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
481 TTGCTGTGGAAGCTCGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
481 TTGCTGTGGAAGCTCGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
601 TGGTTCGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
601 TGGTTCGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
661 CGCGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
661 CGCGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTCAACAAACCATGCAAAATGCTGA 780
721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTCAACAAACCATGCAAAATGCTGA 780
781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAACCGATCAGATGGCGCTGGGCGCAA 840
781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAACCGATCAGATGGCGCTGGGCGCAA 840
841 TCGCGCCCATTAACGAGTCCGGCTCGCGTTGGTTCGCGATATCTCGGTAGTGGGATAGC 900
841 TCGCGCCCATTAACGAGTCCGGCTCGCGTTGGTTCGCGATATCTCGGTAGTGGGATAGC 900
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAAGGATTTTC 960
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAAGGATTTTC 960
961 GCCTGCTGGGCAAAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
961 GCCTGCTGGGCAAAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
1021 AGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAGAAAACCAACCTGGCGCCCAATA 1080
1021 AGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAGAAAACCAACCTGGCGCCCAATA 1080
1081 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGCACTGGCAGCAGCAGGTTT 1140
|||||

1081 CGCAACCGCCTCTCCCGCGGCTTGGCCGATTCAATTAATGCACTGGCAGCAGCAGGTTT 1140
1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGACCAATGCTTCTGCGG 1260
1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGACCAATGCTTCTGCGG 1260
1261 TCAGGACGCCATCGGAAGCTGTGGTGTGAGCTGTGAGGTGAGGTGAGGTGAGGTGAGGT 1320
1261 TCAGGACGCCATCGGAAGCTGTGGTGTGAGCTGTGAGGTGAGGTGAGGTGAGGTGAGGT 1320
1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGAATATGTTTTTGGCGCGAATCAATCAACGTT 1380
1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGAATATGTTTTTGGCGCGAATCAATCAACGTT 1380
1381 CTGGCAATATCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
1381 CTGGCAATATCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTCA 1500
1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTCA 1500
1501 GCCTTCAACCAAGGACCATAGATTAATGAAAACCTGAAGAGGTAATCTGTAATCTGG 1560
1501 GCCTTCAACCAAGGACCATAGATTAATGAAAACCTGAAGAGGTAATCTGTAATCTGG 1560
1561 ATTAACCGCGATTAAGGCTATTAACGGTCTCGCTGAGTCCGTTAGGATTAAGGATTAAG 1620
1561 ATTAACCGCGATTAAGGCTATTAACGGTCTCGCTGAGTCCGTTAGGATTAAGGATTAAG 1620
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAATCTGGAAGAGAAATTCCTCAGGTT 1680
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAATCTGGAAGAGAAATTCCTCAGGTT 1680
1681 CGGCAACTGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
1681 CGGCAACTGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCCGGACAAAGGTTCCAGGACAGCTGTAT 1800
1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCCGGACAAAGGTTCCAGGACAGCTGTAT 1800
1801 CCGTTTACCTCGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCCGTT 1860
1801 CCGTTTACCTCGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCCGTT 1860
1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTACTTCCCTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTCCCTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
2101 GGTCTGACCTTCTCGGTTGACCTGATTAATAACAAACATGATGACAGACACCGATTAC 2160
2101 GGTCTGACCTTCTCGGTTGACCTGATTAATAACAAACATGATGACAGACACCGATTAC 2160
2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220
2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220

```
QY 2221 GCATGGTCAAACATCGACACACGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAAACATCGACACACGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACCGCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACGTGATGAAGGCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACGTGATGAAGGCTG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTAATGGCGCCACCATGGAAGAGCCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTAATGGCGCCACCATGGAAGAGCCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGAGTGTGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGAGTGTGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCGAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCGAATTC 2700

RESULT 10
AR031992 LOCUS AR031992 7475 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5866684.
ACCESSION AR031992
VERSION AR031992.1 GI:5946281
KEYWORDS .
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,F.Xavier.
TITLE Peptidyl inhibitors of viral proteases
JOURNAL Patent: US 5866684-A 1 02-FEB-1999;
FEATURES
source Location/Qualifiers
1..7475
/mol_type="unassigned DNA"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGAACAAACCTTTCCGGGTATGGATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGAACAAACCTTTCCGGGTATGGATGATAGCGCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTATACGATGTCCGAGAGTATCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTATACGATGTCCGAGAGTATCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCCAGCTTCTCGGAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCCAGCTTCTCGGAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGGCTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGGCTGGCAC 240
```

```
Db 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGGCTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCACTCAGTCTGCGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCACTCAGTCTGCGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTCCGACG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTCCGACG 360
QY 361 TGTGTGTGTCGATGTAGAAACGAAAGCGGCTGAAAGCTTGTAAAGCGCGGTCACAATC 420
Db 361 TGTGTGTGTCGATGTAGAAACGAAAGCGGCTGAAAGCTTGTAAAGCGCGGTCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACATAATTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACATAATTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGTGCGCATTTGGGTGTCACGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGCTCGG 660
Db 601 TGTGCGCATTTGGGTGTCACGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGAAGCGGCTGAGTGCCATGTCGCGTTTCAACAAAACGATGCAAAATGCTGA 780
Db 721 CGGAAACGGAAGCGGCTGAGTGCCATGTCGCGTTTCAACAAAACGATGCAAAATGCTGA 780
QY 781 ATGAGGCGCATGCTTCCACATCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 840
Db 781 ATGAGGCGCATGCTTCCACATCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 840
QY 841 TGCAGCGCATTAACGAGTCCGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGCAGCGCATTAACGAGTCCGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACCATACCGAAGACAGCTCATGTTATATCCCGGTTAACCCACCATCAACAGATTTTC 960
Db 901 ACCATACCGAAGACAGCTCATGTTATATCCCGGTTAACCCACCATCAACAGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCGAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCGAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAGAAAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAGAAAACCCCTGGCGCCCAATA 1080
QY 1081 CGAAACCGGCTCTCCCGCGGTTGCGCGATTTATTATGACAGCTGGCACGACAGGTTT 1140
Db 1081 CGAAACCGGCTCTCCCGCGGTTGCGCGATTTATTATGACAGCTGGCACGACAGGTTT 1140
QY 1141 CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG 1200
Db 1141 CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG 1200
QY 1201 GCACAAATCTCATGTTTGAAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGAGCCATCGGAAGCTGTGATGCTGTGAGGTCGTAATAATCACTGCAATAATTCG 1320
Db 1261 TCAGGCGAGCCATCGGAAGCTGTGATGCTGTGAGGTCGTAATAATCACTGCAATAATTCG 1320
```

QY 1321 TGTGCTCAAGCGCACTCCCGTTCGTGATAATGTTTTTGGCCGACATCATACGGTT 1380
DB |||||
QY 1321 TGTGCTCAAGCGCACTCCCGTTCGTGATAATGTTTTTGGCCGACATCATACGGTT 1380
DB |||||
QY 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
DB |||||
QY 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
DB |||||
QY 1441 ATTGTGAGGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTGTTTTCAG 1500
DB |||||
QY 1441 ATTGTGAGGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTGTTTTCAG 1500
DB |||||
QY 1501 GCATCTCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAATCTGTAATCTGG 1560
DB |||||
QY 1501 GCATCTCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAATCTGTAATCTGG 1560
DB |||||
QY 1561 ATTAACGGGATAAAGGCTATTAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT 1620
DB |||||
QY 1561 ATTAACGGGATAAAGGCTATTAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT 1620
DB |||||
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAATCGGAAGAGAAATTCACACAGTT 1680
DB |||||
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAATCGGAAGAGAAATTCACACAGTT 1680
DB |||||
QY 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB |||||
QY 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB |||||
QY 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCAGACAAAGCTGTAT 1800
DB |||||
QY 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCAGACAAAGCTGTAT 1800
DB |||||
QY 1801 CGTTTTACTCGGATGCGGTAGCTTACAAAGCAAGCTGATTGCTTACCCGATCGCTGT 1860
DB |||||
QY 1801 CGTTTTACTCGGATGCGGTAGCTTACAAAGCAAGCTGATTGCTTACCCGATCGCTGT 1860
DB |||||
QY 1861 GAAGCGTTATCGCTGATTTATAAAGAGATCTGCTGCCGAACCCCGCAAAAACCTGGGAA 1920
DB |||||
QY 1861 GAAGCGTTATCGCTGATTTATAAAGAGATCTGCTGCCGAACCCCGCAAAAACCTGGGAA 1920
DB |||||
QY 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1981 CTGCAAGAACCGTACTTCACTCGCGCGCTGATTGCTGCTGACGGGGTTTATCGGTTCAAG 2040
DB |||||
QY 1981 CTGCAAGAACCGTACTTCACTCGCGCGCTGATTGCTGCTGACGGGGTTTATCGGTTCAAG 2040
DB |||||
QY 2041 TATGAAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2041 TATGAAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
DB |||||
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
DB |||||
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
DB |||||
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
DB |||||
QY 2221 GCATGCTCAACATCGACACCAAGCAAGTGAATTTATGTTGTAACGGTACTGCGGACCTTC 2280
DB |||||
QY 2221 GCATGCTCAACATCGACACCAAGCAAGTGAATTTATGTTGTAACGGTACTGCGGACCTTC 2280
DB |||||
QY 2281 AAGGTCACCAATCCAAACCGTTGCTGGCTGCTGAGCGCAGGATTAATAACCGCGCAGT 2340
DB |||||
QY 2281 AAGGTCACCAATCCAAACCGTTGCTGGCTGCTGAGCGCAGGATTAATAACCGCGCAGT 2340
DB |||||
QY 2341 CCGAACNAGAGCTGGCAAAAGAGTTCTTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
DB |||||
QY 2341 CCGAACNAGAGCTGGCAAAAGAGTTCTTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
DB |||||

QY 2401 GAAGCGGTTAATAAAGCAAAACCGCTGGTCCGTAGCGCTGAAGTCTTTACGAGGAAG 2460
DB |||||
QY 2401 GAAGCGGTTAATAAAGCAAAACCGCTGGTCCGTAGCGCTGAAGTCTTTACGAGGAAG 2460
DB |||||
QY 2461 TTGGGAAAGATCCACGTTATTCGCCACCATGGAACGCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2461 TTGGGAAAGATCCACGTTATTCGCCACCATGGAACGCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGTCGCTGCTGCTGCTGCTGCTGCTG 2580
DB |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGTCGCTGCTGCTGCTGCTGCTGCTG 2580
DB |||||
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGAGCGCGAGACTAAATTCGAGCTCG 2640
DB |||||
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGAGCGCGAGACTAAATTCGAGCTCG 2640
DB |||||
QY 2641 AACAAACAACAATAACAATAACAACAACTCGGATCGAGGGAAGATTTCAGAAATTC 2700
DB |||||
QY 2641 AACAAACAACAATAACAATAACAACAACTCGGATCGAGGGAAGATTTCAGAAATTC 2700
DB |||||

RESULT 11
AR207294
LOCUS AR207294 7475 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6372883.
ACCESSION AR207294
VERSION AR207294.1 GI:21506162
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,P.Xavier.
TITLE Antiviral medicaments
JOURNAL Patent: US 6372883-A 1 16-APR-2002;
FEATURES
source Location/Qualifiers
1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGCGCATGATAGCGCCGGAGAGA 60
DB |||||
QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGCGCATGATAGCGCCGGAGAGA 60
DB |||||
QY 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAAACGTTATACGATGTCGACAGTATGCCG 120
DB |||||
QY 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAAACGTTATACGATGTCGACAGTATGCCG 120
DB |||||
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAA 180
DB |||||
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAA 180
DB |||||
QY 181 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTAACATCCCAACCGCTGGCAC 240
DB |||||
QY 181 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTAACATCCCAACCGCTGGCAC 240
DB |||||
QY 241 AACAACTCGCGGCAACACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB |||||
QY 241 AACAACTCGCGGCAACACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB |||||
QY 301 ACGCCCGTCGCAAAATTCGTCGGCGGATTAATCTCGCCCGATCAATCGGGTGCAGCG 360
DB |||||
QY 301 ACGCCCGTCGCAAAATTCGTCGGCGGATTAATCTCGCCCGATCAATCGGGTGCAGCG 360
DB |||||
QY 361 TGGTGGTTCGATGTTAGAACGAAAGCGCGTTCGAAAGCGCTGTAAAGCGCGGTGCAATTC 420
DB |||||
QY 361 TGGTGGTTCGATGTTAGAACGAAAGCGCGTTCGAAAGCGCTGTAAAGCGCGGTGCAATTC 420
DB |||||

QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTGTACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTGTACCCAGA 540
QY 541 CACCCATCAACAGTATATTTCTCCCATGAAGACGGTACGGCACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATATTTCTCCCATGAAGACGGTACGGCACTGGCGGTGGAGCATC 600
QY 601 TGGTCCGATTTGGGTCAACAGCAAACTCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCCGATTTGGGTCAACAGCAAACTCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGAGGCGACTGGAGTGCCATGTCGGTCTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGAGGCGACTGGAGTGCCATGTCGGTCTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGGCAA 840
Db 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGGCAA 840
QY 841 TGGCGGCCATTACCGAGTCCGGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACACCATCAAAACAGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACACCATCAAAACAGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGCGCGCTCTCACTGTTGAAAGAAACACCACTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCGCTCTCACTGTTGAAAGAAACACCACTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAAATGCACTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAAATGCACTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTGACAGCTTATCATGCTGACCGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTCATGTTTGACAGCTTATCATGCTGACCGTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCAATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATAATGTTTTTGGCCGACATCAACCGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATAATGTTTTTGGCCGACATCAACCGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGGA 1440
QY 1441 ATTGTAGCGGATAAACAAATTTCAACAGGAAACAGCCAGTCCGTTTATAGGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAAACAAATTTCAACAGGAAACAGCCAGTCCGTTTATAGGTGTTTTCACGA 1500

QY 1501 GCATTTCCACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG 1560
Db 1501 GCATTTCCACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG 1560
QY 1561 ATTAAACGGCGATAAAGGCTATAACGGTCTCGTCAAGTCCGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGCGATAAAGGCTATAACGGTCTCGTCAAGTCCGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCTCCACAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAACAGCCCGCTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAACAGCCCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTAAACGCAAGCTGATTGCTTACCCTGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTAAACGCAAGCTGATTGCTTACCCTGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTCTGCGGAAACCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTCTGCGGAAACCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAAACCGCTTCACTGCGCGCTGATTGCTGCTGACCGGGGTTATGGTTCAAG 2040
Db 1981 CTGCAAAACCGCTTCACTGCGCGCTGATTGCTGCTGACCGGGGTTATGGTTCAAG 2040
QY 2041 TATGAAAAACCGCAAGTACGACATTAAAGACGTCGGCGTGATTAACGTCGCGGAAGCG 2100
Db 2041 TATGAAAAACCGCAAGTACGACATTAAAGACGTCGGCGTGATTAACGTCGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGATCTGATTAATAAACAACAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGATCTGATTAATAAACAACAACAATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAAGCTGCTCTTAAAGGCGAAACAGCGATGACCATCAACGCGCGGTG 2220
Db 2161 TCCATCGCAAGCTGCTCTTAAAGGCGAAACAGCGATGACCATCAACGCGCGGTG 2220
QY 2221 GCATGGTCCAAACATCGACACACGCAAAAGTAAATATGGTGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACACGCAAAAGTAAATATGGTGTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAAACCGTTGCTGCTGAGCGCAGGTATTAACGCGCGCAGT 2340
Db 2281 AAGGGTCAACCATCCAAAACCGTTGCTGCTGAGCGCAGGTATTAACGCGCGCAGT 2340
QY 2341 CCGAAACAAAGAGCTGCGCAAAAGGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAAACAAAGAGCTGCGCAAAAGGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCCGAAAGATCCAGTATTCGCGCCACCATGGAANAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCCGAAAGATCCAGTATTCGCGCCACCATGGAANAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAAACATCCCGCAGATGTCGCGTTTCTGGTATGCGGTGCGTACTCGCGTGTATCAACGCC 2580
Db 2521 CCGAAACATCCCGCAGATGTCGCGTTTCTGGTATGCGGTGCGTACTCGCGTGTATCAACGCC 2580
QY 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCCGACAGCTAAATTCGAGCTCG 2640

Db	2581		GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAGACGCGCAGACTAATTCAGCTCG	2640
Qy	2641		AACAACAAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGATTTTCAGAATTC	2700
Db	2641		AACAACAAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGATTTTCAGAATTC	2700
RESULT 12				
AX377532				
LOCUS	AX377532	9191 bp	DNA	linear
DEFINITION	Sequence 9 from Patent WO0212553.			PAT 18-MAR-2002
ACCESSION	AX377532			
VERSION	AX377532.1	GI:19573718		
KEYWORDS				
SOURCE	synthetic construct			
ORGANISM	synthetic construct			
REFERENCE	1			
AUTHORS	Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensorf,H. and Muth,J.			
TITLE	Method for detecting mutations in nucleotide sequences			
JOURNAL	Patent: WO 0212553-A 9 14-PEB-2002;			
FEATURES	Nanogen Recognomics GmbH (DE)			
source	Location/Qualifiers			
	1..9191			
	/organism="synthetic construct"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:32630"			
	/note="pMALc2x-muts Plasmid"			
ORIGIN				
Query Match	81.8%;	Score 2698;	DB 6;	Length 9191;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 2701;	Conservative	0;	Mismatches	5;
			Indels	0;
			Gaps	0;
Qy	1	CCGACACCATCGAATGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGA	60	
Db	1	CCGACACCATCGAATGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGA	60	
Qy	61	GTCAATTTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGATGTGCGAGAGTATGCCG	120	
Db	61	GTCAATTTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGATGTGCGAGAGTATGCCG	120	
Qy	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGTAACCGAGTAAACGTTTCTCGCAAAA	180	
Db	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGTAACCGAGTAAACGTTTCTCGCAAAA	180	
Qy	181	CGCGGGAAAAAGTGGAAAGCGCGATGCGCGAGCTGAATTATCATTTCCCAACCGTGGCAC	240	
Db	181	CGCGGGAAAAAGTGGAAAGCGCGATGCGCGAGCTGAATTATCATTTCCCAACCGTGGCAC	240	
Qy	241	AACAACTGGCGGGCAACAGTCGTTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300	
Db	241	AACAACTGGCGGGCAACAGTCGTTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300	
Qy	301	ACGCGCGTTCGCAAAATTCGCGCGGATTAATCTCGCGCGATCACTGGTGGCCAGCG	360	
Db	301	ACGCGCGTTCGCAAAATTCGCGCGGATTAATCTCGCGCGATCACTGGTGGCCAGCG	360	
Qy	361	TGTTGGTGTTCGATGGTGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC	420	
Db	361	TGTTGGTGTTCGATGGTGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC	420	
Qy	421	TTCTCGGCAACGGTCACTGGGCTGATCATTAATCTCCGCTGGATGACAGAGTGCCA	480	
Db	421	TTCTCGGCAACGGTCACTGGGCTGATCATTAATCTCCGCTGGATGACAGAGTGCCA	480	
Qy	481	TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGGTTATTTCTTGATGCTCTCACCAGA	540	
Db	481	TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGGTTATTTCTTGATGCTCTCACCAGA	540	
Qy	541	CACCCATCAACAGTATTATTTTCTCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC	600	

Db 1621 ACCGGAATTAAGTACACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCGCTGTGGCTGAAATCAACCCCGGACAAAGCGTTTCCAGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCGCTGTGGCTGAAATCAACCCCGGACAAAGCGTTTCCAGACAAGCTGTAT 1800
Qy 1801 CGGTTTACCTGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCGCATCGCTGTT 1860
Db 1801 CGGTTTACCTGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCGCATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAGAACTGAAAGCAAGTAAAGCGTAAAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAGAACTGAAAGCAAGTAAAGCGTAAAGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTCGCGCTGATTTGCTGCTGACGGGGTTATCGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTCGCGCTGATTTGCTGCTGACGGGGTTATCGCTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGGATAAGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGGATAAGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCCGTTGACTGATTTAAACAAACATCAATGAATGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCCGTTGACTGATTTAAACAAACATCAATGAATGACACCGGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAAATAAGCGCAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAAATAAGCGCAACAGCGATGACCATCAACGCGCCGCTGG 2220
Qy 2221 GCATGGTCCAAATCGACACACGCAAGTGAATATGCTGTAACGCTGACCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGCAAGTGAATATGCTGTAACGCTGACCGACCTTC 2280
Qy 2281 AAGGTCACCAATCCAAACCGTTCGTTGGCGTGTGACGCGAGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTCGTTGGCGTGTGACGCGAGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2460
Qy 2461 TTGCGGAAAGATCCACGCTATTGCGCGCACCATGAAACCGCCAGAAAGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGCTATTGCGCGCACCATGAAACCGCCAGAAAGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGCTGGTGGTGGTGGTGGTGGTGGTGGT 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGCTGGTGGTGGTGGTGGTGGTGGTGGT 2580
Qy 2581 GCCAGCGCTGTCAGACTGTCGATGAAGCTTGAAGACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGCTGTCAGACTGTCGATGAAGCTTGAAGACCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACAAACAAATAAACAATCAACAACTCGGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATAAACAATCAACAACTCGGGATCGAGGAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCC 2706
Db 2701 GGATCC 2706

RESULT 13

AF097412 6724 bp DNA circular SYN 28-JAN-2002
LOCUS Expression vector pMBP-parallel1, complete sequence.
DEFINITION AF097412
ACCESSION AF097412
VERSION AF097412.1 GI:3983120
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source
1. 6724
/organism="Expression vector pMBP-parallel1"
/mol_type="other DNA"
/db_xref="taxon:84479"
204..1163
/gene="lacIq"
204..1163
/gene="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAC83812.1"
/db_xref="GI:3983121"
/translation="MAELNYIPNRVAQAGKQSLIGVATSSLAHAPSQIVAAIKS
RADLGASVVSVMSVERSGEACKAAVHNLQAQVSGLIINYPLDDQDAIAYVAACTNV
PALFLDVSQTPINSIIFSGEDGTRGLVHVALGHQIALLAGPLSSVARSRLAGW
HXYLRNQIQPIAEREGDSMAGSGFOOTQMLNEGIVPTAMLVANDQMALGARVATE
SGLRVGADISVVGDDTSSCVIPPLTTIKODFRLLGOTSVDRLQLLSQGOAVKGNQ
LLPVSLVKRKTTLAPNTQTASPRALADSLMQLARQVSRLESQ"
1453..1481
/note="tac; inducible promoter"
1528..2994
/gene="male"
1528..2994
/gene="male"
/codon_start=1
/transl_table=11
/product="maltose binding protein"
/protein_id="AAC83813.1"
/db_xref="GI:3983122"
/translation="MKIEGKLVIMWNGKGYNGLAEVGKFKEDTGIKVTVEHPDKL
BEKFPVAATGDPDIIFWAHDRCFGYAGSGLLAETPDKAFQDKLIYPTWDAVRNG
KLIAVPIAVEALSLLYKNDLLPNPKTWEIIPALKEKAKGKSALMFLNQBPYFTWP
LIAAGGVAFYENGKIDIKQVGNAGAKAGLTFVDLIKNKHNNADTDYSIAEAAFL
NKGEMTINGPWAWNSIDTSKVYGVTVLPTFKGQSPKFPVGLSAGINAAKPNKEL
AKEFLNLLTDEGLEAVNKDKPLGVALKSYEEELAKDPRIATAWNAQKEIMPNL
PQMSAFWAVRTAVNAASGQTVDEALKDAQTNSSNNNNNNNNNNLNGIDITENLYF
QCAMDPFAGLRRAQLVRPLSNLRPAVSRHAPVSLALAVLQRRDWNPGVTLQNLRL
AAHPFPASWRNSEARTORPSQSLRSLNGEWQLGCGFGG"
2633..2688
/gene="male"
/note="protein spacer region"
2689..2709
/gene="male"
/note="Tobacco etch virus protease cleavage site"
2711..2810
/gene="male"
misc_feature
misc_feature
misc_feature


```
gene /note="multiple cloning site"
3426..4356
/gene="bla"
promoter /gene="bla"
3426..3454
/gene="bla"
CDS 3496..4356
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAC83814.1"
/db_xref="GI:3983123"
/translation="MSIQHPRVALIPPEAAFLPVFAHPETLVKVKDAEDQLGARVGY
IELDNSKILIEFPERESPMSFTKDLGLAVLSRVDDAGEQGRIRHYNQDNLVE
YSPVTEKHLTDGMTVRELCSAAITWSDNTAANLLLTITGGPKELTAFIHNMGDHLRL
DRWPELNEAIPNDRDITMPVAMATLLRLTIGELLTLASRQOLI DMWEADKRVGRL
LRSLPAGWFTADKSGAGERSGRIIAALGPDGKPSRIWVIYTTGSQATMDERNQIA
EIGASLIKHW"
```

ORIGIN

```
Query Match 81.1%; Score 2676.2; DB 12; Length 6724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60
DB 1 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60

QY 61 GTCAATTACAGGTGTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCG 120
DB 61 GTCAATTACAGGTGTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCG 120

QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGTGAACACGAGCCAGCCAGCTTTCTGGGAAA 180
DB 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGTGAACACGAGCCAGCCAGCTTTCTGGGAAA 180

QY 181 CGCGGAAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATTCGCCAACCGCTGGGC 240
DB 181 CGCGGAAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATTCGCCAACCGCTGGGC 240

QY 241 AACAACTGCGGGGAAAACAGTGTGCTGATTTGGCGTGTGCACCTCCAGCTCGGCCCTGC 300
DB 241 AACAACTGCGGGGAAAACAGTGTGCTGATTTGGCGTGTGCACCTCCAGCTCGGCCCTGC 300

QY 301 ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360

QY 361 TGGTGGTTCGATGGTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGTGCACAATC 420
DB 361 TGGTGGTTCGATGGTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGTGCACAATC 420

QY 421 TTCTCGCGCAACGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGGATGCCA 480
DB 421 TTCTCGCGCAACGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGGATGCCA 480

QY 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTGATGTCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTGATGTCTCTGACCCAGA 540

QY 541 CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

QY 601 TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG 660
DB 601 TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG 660

QY 661 CGCGTCTGGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAACGGGAAGGCGACTGGAGTGCAATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
```

```
DB 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGTGTGGTGGCAACGATCAGATGCGCTGGGCGCAA 840
DB 781 ATGAGGGCATCGTTCCCACTCGCATGTGTGGTGGCAACGATCAGATGCGCTGGGCGCAA 840
QY 841 TGGCGGCATTTACAGAGTCCGGGCTCGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
DB 841 TGGCGGCATTTACAGAGTCCGGGCTCGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACACCATCAACACGAGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACACCATCAACACGAGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACACGAGCGTGGACCGCTTCTCTCAGGCGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACACGAGCGTGGACCGCTTCTCTCAGGCGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCCTCTCACTGGTGAAGAAAAACCACTGGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTGGCCCTCTCACTGGTGAAGAAAAACCACTGGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGCACGACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGCACGACAGGTTT 1140
QY 1141 CCCGACTCGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 1141 CCCGACTCGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGACCGGTGCACCAATGCTTCGGCG 1260
DB 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGACCGGTGCACCAATGCTTCGGCG 1260
QY 1261 TCAGGCAGCATCGGAAGCTGTGTTGGCTGTCAGCTGCTAATCACTGCATTAATTCG 1320
DB 1261 TCAGGCAGCATCGGAAGCTGTGTTGGCTGTCAGCTGCTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACCTCCCGTCTCGATTAATGTTTTGCGCGGACATCATACGGTT 1380
DB 1321 TGTGCTCAAGCGGCACCTCCCGTCTCGATTAATGTTTTGCGCGGACATCATACGGTT 1380
QY 1381 CTGGCAAAATATTTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
DB 1381 CTGGCAAAATATTTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGTAACAAATTTTCAACAGGAACAGCCAGTCCGTTTGGTGTTCACGA 1500
DB 1441 ATTGTGAGCGGTAACAAATTTTCAACAGGAACAGCCAGTCCGTTTGGTGTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAAGAGGTAAATCTGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAAGAGGTAAATCTGTAATCTGG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAGCTGAT 1800
DB 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAGCTGAT 1800
QY 1801 CCGTTTACTCTGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCCGATCCCTGTT 1860
DB 1801 CCGTTTACTCTGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCCGATCCCTGTT 1860
```



```
Db      841  TCCGCGCCATTACCGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy      901  ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACCAACCAATCAAAACAGGATTTTC 960
Db      901  ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACCAACCAATCAAAACAGGATTTTC 960
Qy      961  GCCTGCTGGGGCAAAACCGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db      961  GCCTGCTGGGGCAAAACCGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Qy     1021  AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAACCAACCTGGCGGCCCAATA 1080
Db     1021  AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAACCAACCTGGCGGCCCAATA 1080
Qy     1081  CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCAATTAATGCACTGGCGACACAGGTTT 1140
Db     1081  CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCAATTAATGCACTGGCGACACAGGTTT 1140
Qy     1141  CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db     1141  CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Qy     1201  GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGGTCACCAATGCTTCTGGCG 1260
Db     1201  GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGGTCACCAATGCTTCTGGCG 1260
Qy     1261  TCAGCAGCCATCGGAAGCTGTGGTATGCTGTGCAGGTCGTAATCACTGCATAAATTCG 1320
Db     1261  TCAGCAGCCATCGGAAGCTGTGGTATGCTGTGCAGGTCGTAATCACTGCATAAATTCG 1320
Qy     1321  TGTGCTCAAGCGCACTCCCGTCTGATATGTTTTTGGCGCGACATCATACAGGTT 1380
Db     1321  TGTGCTCAAGCGCACTCCCGTCTGATATGTTTTTGGCGCGACATCATACAGGTT 1380
Qy     1381  CTGGCAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db     1381  CTGGCAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Qy     1441  ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGCTTTAGGTGTTTTTCAGGA 1500
Db     1441  ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGCTTTAGGTGTTTTTCAGGA 1500
Qy     1501  GCATTTCCAAACAGGACCATAGATTATGAAAACTGAAGAAAGGTAATACTGGTAAATCTGG 1560
Db     1501  GCATTTCCAAACAGGACCATAGATTATGAAAACTGAAGAAAGGTAATACTGGTAAATCTGG 1560
Qy     1561  ATTTACGGCGATAAAGGCTATAACGCTCTCGCTGAGTCGTAAGAAATTCAGAAAGAT 1620
Db     1561  ATTTACGGCGATAAAGGCTATAACGCTCTCGCTGAGTCGTAAGAAATTCAGAAAGAT 1620
Qy     1621  ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT 1680
Db     1621  ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT 1680
Qy     1681  GCGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCAACAGCGCTTTGGTGCTAC 1740
Db     1681  GCGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCAACAGCGCTTTGGTGCTAC 1740
Qy     1741  GCTCAATCTGGCGCTGTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db     1741  GCTCAATCTGGCGCTGTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy     1801  CCGTTTACCTGGGATGCGCTGACGTTACAAACGGCAAGCTGATTCCTTACCCGATCGCTGT 1860
Db     1801  CCGTTTACCTGGGATGCGCTGACGTTACAAACGGCAAGCTGATTCCTTACCCGATCGCTGT 1860
Qy     1861  GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTCGGAACCCGCCAAAACCTGGGAA 1920
Db     1861  GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTCGGAACCCGCCAAAACCTGGGAA 1920
Qy     1921  GAGATCCCGCGCTGATATAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db     1921  GAGATCCCGCGCTGATATAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
```

```
Qy      1981  CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db      1981  CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy     2041  TATGAAACGCGCAAGTAGTACGACATTAAGACCTGGGCGTGGATACGCTGGCGCGAAACGG 2100
Db     2041  TATGAAACGCGCAAGTAGTACGACATTAAGACCTGGGCGTGGATACGCTGGCGCGAAACGG 2100
Qy     2101  GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAAACAACAACATGAATGCAGACACCGATTAC 2160
Db     2101  GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAAACAACAACATGAATGCAGACACCGATTAC 2160
Qy     2161  TCCATCGAGAAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db     2161  TCCATCGAGAAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy     2221  GCATGGTCCAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
Db     2221  GCATGGTCCAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
Qy     2281  AAGGTCACACCATCCAAACCGTTCTGGCGCTGCTGAGCGAGGATTAACGCCGCCAGT 2340
Db     2281  AAGGTCACACCATCCAAACCGTTCTGGCGCTGCTGAGCGAGGATTAACGCCGCCAGT 2340
Qy     2341  CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG 2400
Db     2341  CCGAACAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG 2400
Qy     2401  GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAACTCTTACGAGGAAGAG 2460
Db     2401  GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAACTCTTACGAGGAAGAG 2460
Qy     2461  TTGCGAAGAGTCCACGCTATTGCGGCCACCACTGAAAAACGCCAGAAAGGTGAATCATG 2520
Db     2461  TTGCGAAGAGTCCACGCTATTGCGGCCACCACTGAAAAACGCCAGAAAGGTGAATCATG 2520
Qy     2521  CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db     2521  CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Qy     2581  GCCAGCGTCTGACACTGCTGATGAAGCCCTGAAAGACGCGACACTAAATTCGAGCTCG 2640
Db     2581  GCCAGCGTCTGACACTGCTGATGAAGCCCTGAAAGACGCGACACTAAATTCGAGCTCG 2640
```

```
RESULT 15
AF031088      6706 bp      DNA      circular SYN 07-MAY-2001
LOCUS      Shuttle vector pMAL-pIII, complete sequence.
DEFINITION  Shuttle vector pMAL-pIII
ACCESSION  AF031088
VERSION    AF031088.1 GI:2623821
KEYWORDS  Shuttle vector pMAL-pIII
SOURCE     Shuttle vector pMAL-pIII
ORGANISM   Shuttle vector pMAL-pIII
other sequences; artificial sequences; vectors.
REFERENCE  1 (bases 1 to 6706)
AUTHORS    Zwick,M.B., Bonnycastle,L.L., Noren,K.A., Venturini,S., Leong,E.,
Barbas,C.F. III, Noren,C.J. and Scott,J.K.
The maltose-binding protein as a scaffold for monovalent display of
peptides derived from phage libraries
Anal. Biochem. 264 (1), 87-97 (1998)
99002881
9784192
PUBMED
2 (bases 1 to 6706)
Noren,K.A. and Noren,C.J.
Direct Submission
Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road,
Beverly, MA 01915, USA
Location/Qualifiers
1..6706
/organism="Shuttle vector pMAL-pIII"
/mol_type="genomic DNA"
```

```

/db_xref="taxon:69327"
/notes="derived from commercially available vector pMal-p2
(New England Biolabs)"
15..1163
/gene="lacI"
/promoter
15..46
/notes="E. coli gene"
81..1163
/gene="lacI"
/notes="transcriptional repressor"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAB86558.1"
/db_xref="GI:2623822"
/translation="MKPTLVDAEYAGVSVQTVSRVNVQASHVSATKREVEAMAE
LNYINRVAQQLAGKSLIGVATSSLAHAPSQIVAAIKSRADQLGASVVSVER
GVEACKAVHNLLAQRVSLGIINYPDQDDAIAVEAACTNPALFLDVSQDTPINSII
FSHEDGTRLGVEHLVALGHQIALLAGLSVSLRGLAGWHKYLTRNQIPIABREG
DMSMSGFQTMQLNEGIVPTAMLVANDQMALGAMRAITSSGLRVGADIISVGVYDIT
EDSSCYIPPLTTIKQDFRLLGQTSVDRLLQLSQGAVKGNQLLPVSLVKRKTLLAPNT
QTASPRALADSLMLQARQVSRLESQ"
1406..1433
/notes="lac promoter"
1528..2976
/gene="male-lacZa fusion"
/notes="E. coli gene fusion"
1528..2976
/gene="male-lacZa fusion"
/notes="MBP-lacZa; MBP modified to include phage M13
secretion signal sequence"
/codon_start=1
/transl_table=11
/product="maltose binding protein-lacZ alpha peptide
fusion protein precursor"
/protein_id="AAB86559.1"
/db_xref="GI:2623823"
/translation="MKKLLFAIPLVFPVYSHSADMKIEBGLKVIWINGDKYGLAEV
GKPEKDTGIVTVHEDPKIEEPQVAATGDGPDIIFWAHDHFGVQAQGLLAEIIP
DKAFQDKLYPTTVDAVRNGKLIAYPIAVEALSILYKNDLLPNPKTWESIIPALDKEL
KAGKSAIMFNLPQPYFTWPLIIAADGVAFKPYKGVYDIKDVGVDNAGAKGLTFLVD
LKNKHNADPDYSIAEANKGETAMTINGPHAWSNIDTSKNYGVTVLPTFKGQPS
KPFVGVLSAGINAASPNKELAKFLENYLLTDEGLEAVNKDKPLIGVALKSYEEELAK
DPIAATMENAKQIEPNIPTQMSAFYAVRTAVINAASGRQTVDEALDKDAQTNSSN
NNNNNNNNIEGRIIEFGSSRVDLQASLALAVLQRRDWNENPGVTLNRLAAHPFP
ASWNSSEARTDRPSQQLRSLNGEWQLGCFG"
1528..1581
/gene="male-lacZa fusion"
/notes="M13 secretion signal sequence"
1582..2973
/gene="male-lacZa fusion"
/product="maltose binding protein-lacZ alpha peptide
fusion protein"
3478..4338
/gene="bla"
/notes="R-plasmid ampicillin resistance gene"
3478..4338
/gene="bla"
/notes="penicillinase"
/codon_start=1
/transl_table=11
/product="RTEM-1 beta-lactamase"
/protein_id="AAB86560.1"
/db_xref="GI:2623824"
/translation="MSIQHPRVALIPFFAFLPFAVHPETLVKVKDAEDQLGARVGY
IELDLNSKILLESFRRPEERPMSTFKVLLCGVALSRVDAQBOIGRRIRHYSQNDLVE
YSPVTEKHLTDGHTVRELCSAAITMSDNTAANLLTIIGCKELTAFLHMGDHYTRL
DRWPELNEAIPINDERDTTTPMVAWMTLRKLLTGELLTLASRQQLIDWNEADKVAQPL
LRSLPAGWFIADKSGNGERSGIITAAIGPDGKPSRIVVIYITGSGATMDERNRQIA
EIGASLIKHH"
4381..4893
/standard_name="M13 origin (+ -)"

```

ORIGIN

```

Query Match      79.7%; Score 2631; DB 12; Length 6706;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 60; Gaps 1;

QY 1  CCGACACCATCGAATGTGTGAACACAGTAAACCTTTCCGGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 1  CCGACACCATCGAATGTGTGAACACCTTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60

QY 61  GTCAATTCAGGGTGGTCAATGTGTGAACACAGTAACTGTATACGATGTGCGAGAGTATCCG 120
Db 61  GTCAATTCAGGGTGGTCAATGTGTGAACACAGTAACTGTATACGATGTGCGAGAGTATCCG 120

QY 121  GTGTCTCTTATACAGACCGTTTCCCGGTGTGTGAACACAGCGCAGCCAGCTTTCTGCGGAAAA 180
Db 121  GTGTCTCTTATACAGACCGTTTCCCGGTGTGTGAACACAGCGCAGCCAGCTTTCTGCGGAAAA 180

QY 181  CCGGGGAAAAAGTGGAGACGGCGATGCGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db 181  CCGGGGAAAAAGTGGAGACGGCGATGCGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240

QY 241  AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGGCACTCCAGTCTCGCCCTGC 300
Db 241  AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGGCACTCCAGTCTCGCCCTGC 300

QY 301  ACGGGCGCTCGCAATTTGTCGGCGGATTAATATCTCGCGCGATCACTGGGTGCCAGCG 360
Db 301  ACGGGCGCTCGCAATTTGTCGGCGGATTAATATCTCGCGCGATCACTGGGTGCCAGCG 360

QY 361  TGGTGGTGTGCGATGTTAGAACGAGCGGCTGCAAGCCCTGTAAGCGGGGTGCACAATC 420
Db 361  TGGTGGTGTGCGATGTTAGAACGAGCGGCTGCAAGCCCTGTAAGCGGGGTGCACAATC 420

QY 421  TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
Db 421  TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480

QY 481  TTCTCTGGAAGCTGCTGCACCTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACAGA 540
Db 481  TTCTCTGGAAGCTGCTGCACCTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACAGA 540

QY 541  CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGGACTGGCGCTGGAGCATC 600
Db 541  CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGGACTGGCGCTGGAGCATC 600

QY 601  TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCG 660
Db 601  TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCG 660

QY 661  CGCGCTCTGCTCTGGCTGCGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661  CGCGCTCTGCTCTGGCTGCGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721  CGGAACGGGAAGCGCATGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721  CGGAACGGGAAGCGCATGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780

QY 781  ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGATGAGCTGGCGTGGGCGCAA 840
Db 781  ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGATGAGCTGGCGTGGGCGCAA 840

QY 841  TCGCGGCCATTACCGAGTCCGGGCTGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841  TCGCGGCCATTACCGAGTCCGGGCTGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATAG 900

QY 901  ACGATACCGAAGACGCTCATGTTATATCCCGCGCTTAAACCAATCAAAACAGGATTTTC 960
Db 901  ACGATACCGAAGACGCTCATGTTATATCCCGCGCTTAAACCAATCAAAACAGGATTTTC 960

QY 961  GCCTCTGGGGCAAAACAGCGTGGACCGCTTGCTCAACTCTCTCAGGGCCAGGGGTGA 1020
Db 961  GCCTCTGGGGCAAAACAGCGTGGACCGCTTGCTCAACTCTCTCAGGGCCAGGGGTGA 1020

```

QY 1021 AGGGCAATCAGCTGTGTGCCGCTCTCACTGGTGAAAAAAGAAAAACCAACCTGTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTGCCGCTCTCACTGGTGAAAAAAGAAAAACCAACCTGTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACAGCTGGCAGACAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACAGCTGGCAGACAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTATGTTTGAACAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTATGTTTGAACAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGCTGAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGCTGAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATGAGCTGTTGACATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGACATTAATCATCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTTGAGCGGATACAAATTTTACACAGAAACAGCCAGTCCGTTAGTGTTCAGGA 1500
Db 1441 ATTTGAGCGGATACAAATTTTACACAGAAACAGCCAGTCCGTTAGTGTTCAGGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAG----- 1524
Db 1501 GCACCTTCAACCAAGGACCATAGCATATGAAAAATTAATTCGCAATTCCTTAGTG 1560
QY 1525 -----ATTATGAAAACTGGAAGAGTAAACTGGTAATCTGG 1560
Db 1561 GTACCTTTCTATTCTCACTCGGCCGATATGAAATCGAAGAGTAAACTGGTAATCTGG 1620
QY 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGT 1620
Db 1621 ATTAACGGGATAAAGGCTATAACGGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGT 1680
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGTT 1680
Db 1681 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGTT 1740
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1741 GCGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGCGCTTTGGTGGCTAC 1800
QY 1741 GCTCAATCTGGCCTGTGGCTGAATCAACCCGGCAAGAGCTTCAGGACAGCTGTAT 1800
Db 1801 GCTCAATCTGGCCTGTGGCTGAATCAACCCGGCAAGAGCTTCAGGACAGCTGTAT 1860
QY 1801 CCGTTTACCTGGATGCGGTAGCTTACACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1861 CCGTTTACCTGGATGCGGTAGCTTACACGCAAGCTGATTGCTTACCCGATCGCTGT 1920
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAGATCTGCTCCGAAACCCGCAAAAAACCTGGGAA 1920
Db 1921 GAAGCGTTATCGCTGATTTTAAACAAGATCTGCTCCGAAACCCGCAAAAAACCTGGGAA 1980
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTGAAGCGCGCTGATGTTCAAC 1980
Db 1981 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTGAAGCGCGCTGATGTTCAAC 2040
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGACGGGGTTTATCGGTTCAAG 2040
Db 2041 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGACGGGGTTTATCGGTTCAAG 2100

QY 2041 TATGAAAAACGCAAGTAGCAGCATTTAAAGACGTGGCGCTGGATAACGCTGGCGCGGAAACGG 2100
Db 2101 TATGAAAAACGCAAGTAGCAGCATTTAAAGACGTGGCGCTGGATAACGCTGGCGCGGAAACGG 2160
QY 2101 GGTCTGACCTTCTCGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2161 GGTCTGACCTTCTCGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2220
QY 2161 TCCATCGCAGAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGGCCGCTGG 2220
Db 2221 TCCATCGCAGAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGGCCGCTGG 2280
QY 2221 GCATGCTCCACATCGACACGACGAAAGTGAATTAATGGTGAACGGTACTCCGACCTTC 2280
Db 2281 GCATGCTCCACATCGACACGACGAAAGTGAATTAATGGTGAACGGTACTCCGACCTTC 2340
QY 2281 AAGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAAGCGCCCACT 2340
Db 2341 AAGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAAGCGCCCACT 2400
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2401 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2460
QY 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAAG 2460
Db 2461 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAAG 2520
QY 2461 TTGGCGAAAGATCCACGTTATTGGCCGCAACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2521 TTGGCGAAAGATCCACGTTATTGGCCGCAACCATGGAACCGCCAGAAAGGTGAATCATG 2580
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCGTGAATCAACGCC 2580
Db 2581 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCGTGAATCAACGCC 2640
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2641 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2700
QY 2641 AACCAACAAACAAATTAACAATAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2701 AACCAACAAACAAATTAACAATAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2760
QY 2701 GGATCCTCT 2709
Db 2761 GGATCCTCT 2769

Search completed: August 1, 2005, 17:21:49
Job time : 9285 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgacacatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gss1:
9: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	23.6	794	5	BQ751936
2	770.6	23.4	779	7	CK394523
3	768	23.3	769	7	CK118014
4	762	23.1	793	5	BQ751655
5	755.4	22.9	806	5	BQ751220
6	684.8	20.8	689	6	CB863541
7	676.4	20.5	712	1	AL038548
8	676	20.5	710	8	AY080106
9	647.4	19.6	650	1	AL044483
10	641	19.4	641	4	BJ063872
11	615.8	18.7	624	6	CB863814
12	612	18.5	794	1	AL045353
13	603	18.3	752	9	CL655575
14	589.8	17.9	608	8	AY080094
15	586.4	17.8	589	1	AL037742
16	581.2	17.6	601	1	AV594538
17	579	17.5	579	1	AL043868
18	570	17.3	782	7	CK394246
19	569	17.2	774	1	AL045337
20	563.2	17.1	585	9	CL658017
21	562.6	17.0	723	1	AL039416
22	549.4	16.6	753	1	AL045341
23	546.8	16.6	782	1	AL037051
24	545.4	16.5	547	6	CA895956

C 25	545	16.5	648	7	CK394349	CK394349	hgada2D0
C 26	540.2	16.4	569	1	AL039077	AL039077	DKF2p5566G
C 27	534.4	16.2	644	6	CB862524	CB862524	HH04B08u
C 28	534	16.2	734	1	AL048427	AL048427	DKF2p586H
C 29	529	16.0	740	1	AL042909	AL042909	DKF2p434J
C 30	528	16.0	720	6	CA881974	CA881974	K0994C10-
C 31	523.2	15.9	528	1	AL038811	AL038811	DKF2p566O
C 32	519.4	15.7	521	6	CA890032	CA890032	B0136F03-
C 33	519	15.7	607	6	CB862099	CB862099	HH06A02Y
C 34	516.4	15.6	518	6	CA886082	CA886082	B0124F04-
C 35	510.2	15.5	686	1	AL044407	AL044407	DKF2p434D
C 36	509.4	15.4	511	6	CA895436	CA895436	B0192A07-
C 37	498	15.1	498	1	AL039076	AL039076	DKF2p566G
C 38	494.8	15.0	1025	1	AL038025	AL038025	DKF2p566C
C 39	487.6	14.8	498	7	CK394397	CK394397	hgada2H1
C 40	477.6	14.5	491	6	CB862715	CB862715	HH03F14u
C 41	465.8	14.1	679	1	AL039128	AL039128	DKF2p566K
C 42	462.8	14.0	490	1	AL039589	AL039589	DKF2p434D
C 43	458	13.9	458	6	CA887583	CA887583	B0137H04-
C 44	456.4	13.8	471	1	AL039649	AL039649	DKF2p434G
C 45	449.2	13.6	612	6	CB863409	CB863409	HH04K05W

ALIGNMENTS

RESULT 1
BQ751936
LOCUS BQ751936 794 bp mRNA linear EST 18-JUL-2002
DEFINITION EST632499 DSCT Colletotrichum trifolii cDNA clone pDST8-67, mRNA
sequence.
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 794)
AUTHORS Samad,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
Cheung,F. and Fraser,C.M.
TITLE ESTs from mycelia of Colletotrichum trifolii race 1
JOURNAL Unpublished (2002)
COMMENT Other ESTs: EST632498
Contact: Deborah A. Samad
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbys@puccini.crl.umn.edu
TIGR sequence name: MTSAH67TV More information is available at:
www.medicago.org
Seq primer: lgtA AtA CgA Ctc Act AtA ggg C).
Location/Qualifiers
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDST8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2%glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"

FEATURES
source

1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDST8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2%glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 28p2; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gptI from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce

E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

```
Query Match      23.6%; Score 778; DB 5; Length 794;
Best Local Similarity 98.7%; Pred. No. 4.9e-219; Indels 0; Gaps 0;
Matches 784; Conservative 0; Mismatches 10;

Qy 3 GACACCATCGAATCGTGCAAAACCTTTTCGGGTATGGCATATAGCGCCGGAAGAGAGT 62
Db 1 GACACTTTCGAATCGCGCAAAACCTTTTCGGGTATGGCATATAGCGCCGGAAGAGAGT 60
Qy 63 CAATTCAGGTTGGTAATGTGAACACCAAGTAACGTTATACGATGTGCGAGATATGCCGGT 122
Db 61 CAATTCAGGTTGGTAATGTGAACACCAAGTAACGTTATACGATGTGCGAGAGATATGCCGGT 120
Qy 123 GTCTCTTATCAGACCGTTTTCGGCGTGTGTGAACACGAGCCAGCCAGTTTCTGCGAAAACG 182
Db 121 GTCTCTTATCAGACCGTTTTCGGCGTGTGTGAACACGAGCCAGCCAGTTTCTGCGAAAACG 180
Qy 183 CGGAAAAAGTGAAGCGCGGATGGCGAGTGAATTAATTCCTCAACCGCGTGCGCAAA 242
Db 181 CGGAAAAAGTGAAGCGCGGATGGCGAGTGAATTAATTCCTCAACCGCGTGCGCAAA 240
Qy 243 CAACCTGCGGCAACAGTGTCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGCAC 302
Db 241 CAACCTGCGGCAACAGTGTCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGCAC 300
Qy 303 GCGCGCTGCGAAATTTGTCGCGCGGATTAATTCGCGCGGATCAATTCGGGTGCCAGCGT 362
Db 301 GCGCGCTGCGAAATTTGTCGCGCGGATTAATTCGCGCGGATCAATTCGGGTGCCAGCGT 360
Qy 363 GTGTGTGCGATGTGTAAGCAAGCGCGTGAAGCGTGAAGCGCGGTGCAAACTCTT 422
Db 361 GTGTGTGCGATGTGTAAGCAAGCGCGTGAAGCGTGAAGCGCGGTGCAAACTCTT 420
Qy 423 CTCGCGCAACGCGTCAGTGGGCTGATCAATTAACCTATCCGCTGGATGACCGAGTGCATT 482
Db 421 CTCGCGCAACGCGTCAGTGGGCTGATCAATTAACCTATCCGCTGGATGACCGAGTGCATT 480
Qy 483 GCTGTGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACGACAA 542
Db 481 GCTGTGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACGACAA 540
Qy 543 CCGATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTG 602
Db 541 CCGATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTG 600
Qy 603 GTCGATTGGGTACACGCAATCGCGTGTAGCGGCGCATTAAGTCTGTCTCGCG 662
Db 601 GTCGATTGGGTACACGCAATCGCGTGTAGCGGCGCATTAAGTCTGTCTCGCG 660
Qy 663 CGTCTGGTCTGGTGGCTGCGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 661 CGTCTGGTCTGGTGGCTGCGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720
Qy 723 GAACGGGAAGCGACTGAGTGCCATGTCGGGTTTTCAACCAACCATGCAATGCTGAAT 782
Db 721 GAACGGGAAGCGACTGAGTGCCATGTCGGGTTTTCAACCAACCATGCAATGCTGAAT 780
Qy 783 GAGGGCATCGTTCC 796
Db 781 GAGGGCATCGTTCC 794
```

RESULT 2
CK394523 hgada4D08 Gland Cell Amplified cDNA Library Heterodera glycines
LOCUS
DEFINITION

cDNA, mRNA sequence.
CK394523
VERSION CK394523.1 GI:40389794
KEYWORDS EST.
SOURCE Heterodera glycines
ORGANISM Heterodera glycines
REFERENCE 1 Tylenchida; Nematoda; Chromadorea; Tylenchida; Tylenchina; Eukaryota; Metazoa; Heteroderidae; Heteroderinae; Heterodera.
AUTHORS Wang,X., Allen,R., Gao,B., Goellner,M., Maier,T., Baum,T., Hussey,R. and Davis,E.
TITLE Submission: Wang, Allen, Gao, Goellner, Maier, Baum, Hussey, Davis Unpublished (2003)
JOURNAL Contact: Tom Maier
COMMENT Department of Plant Pathology, Baum Lab
Iowa State University
351 Bessey Hall, Ames, IA 50011, USA
Tel: 515-294-8854
Fax: 515-294-9420
Email: trmaier@iastate.edu
Heterodera glycines Gland Cell Amplified cDNA Library, single pass sequence.

FEATURES

Location/Qualifiers
1..779
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell Amplified cDNA Library"
/note="Organ: gland cell; Vector: pSportII"

ORIGIN

```
Query Match      23.4%; Score 770.6; DB 7; Length 779;
Best Local Similarity 99.4%; Pred. No. 7.6e-217; Indels 0; Gaps 0;
Matches 773; Conservative 0; Mismatches 5;

Qy 87 CCAGTAACGTTATACGATGTGCGAGAGTATGCGCGTGTCTCTTATCAGACCGTTTCCCGC 146
Db 779 CCAGTAACGTTATACGATGTGCGAGAGTATGCGCGTGTCTCTTATCAGAGTTTCNCGC 720
Qy 147 GTGTGAACCAAGCGCGACCGCTTTCTGCGAAACCGCGGAAAGTGGAGCGCGATG 206
Db 719 GTGTGAACCAAGCGCGACCGCTTTCTGCGAAACCGCGGAAAGTGGAGCGCGATG 660
Qy 207 GCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACCTGGCGGCAACAGTCTGTTG 266
Db 659 GCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACCTGGCGGCAACAGTCTGTTG 600
Qy 267 CTGATTGGCGTTGCGACCTCCAGTCTGGCCCTGCACGCGCGCTCGCAAAATTTGTCGCGCG 326
Db 599 CTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCGCTCGCAAAATTTGTCGCGCG 540
Qy 327 ATTAATCTCGCGCGATCAACTGGGTGCGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 386
Db 539 ATTAATCTCGCGCGATCAACTGGGTGCGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Qy 387 GCGGTGAAGCTGTAAAGCGCGGTGCAACAATCTTCGCGCAACCGCTCAGTGGGCTG 446
Db 479 GCGGTGAAGCTGTAAAGCGCGGTGCAACAATCTTCGCGCAACCGCTCAGTGGGCTG 420
Qy 447 ATCATTAACTATCCGCTGGATGACCGAGTGCCTATGCTGTGGAAGTGCCTGCACTAAT 506
Db 419 ATCATTAACTATCCGCTGGATGACCGAGTGCCTATGCTGTGGAAGTGCCTGCACTAAT 360
Qy 507 GTTCCGCGTTATTTCTTGATGTCTGTGACGACACCCATCAACAGTATATTTTCTCC 566
Db 359 GTTCCGCGTTATTTCTTGATGTCTGTGACGACACCCATCAACAGTATATTTTCTCC 300
Qy 567 CATGAAGACCGGTACGCGACTGGGCGTGGAGCATCTGGTCTGCTATTTGGTTCACGCAAAATC 626
Db 299 CATGAAGACCGGTACGCGACTGGGCGTGGAGCATCTGGTCTGCTATTTGGTTCACGCAAAATC 240
```


COMMENT

Other ESTs: EST632217
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAF59TV More information is available at:
 www.medicago.org
 Seq primer: (gtA AtA CgA CtC Act AtA ggg C).

FEATURES

source

Location/Qualifiers
 1..793
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDSC76-59"
 /tissue_type="mycelia"
 /dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSC7"

/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl1 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match
 Best Local Similarity 23.1%; Score 762; DB 5; Length 793;
 Matches 777; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
 3 GACACATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCCCGGAGAGAGT 62
 1 GACATCTTCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCCCGGAGAGAGT 60
 63 CAATTCAGGGTGGTAATGTGAACACCGATTAACGATGTCGAGAGTATGCCGGT 122
 61 CAATTCAGGGTGGTAATGTGAACACCGATTAACGATGTCGAGAGTATGCCGGT 120
 123 GTCTCTTATCAGACCGTTTCCCGGTTGGTGAACCGAGCCACGTTTCTGCGAANAACG 182
 121 GTCTCTTATCAGACCGTTTCCCGGTTGGTGAACCGAGCCACGTTTCTGCGAANAACG 180
 183 CGGGAANAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGCACAA 242
 181 CGGGAANAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGCACAA 240
 243 CAACCTGGCGGCAACAGTGTGCTGATTTGGGTTGCCACCTCCAGTTCGGCCCTGCAC 302
 241 CAACCTGGCGGCAACAGTGTGCTGATTTGGGTTGCCACCTCCAGTTCGGCCCTGCAC 300
 303 CGCGCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGGTGGCAGCGTG 362
 301 CGCGCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGGTGGCAGCGTG 360
 363 GTGGTGTGATGTGAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCTGACACATCTT 422
 361 GTGGTGTGATGTGAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCTGACACATCTT 420
 423 CTCGCGCAACCGCTCAGTGGGCTGATCACTTAATCTCCGCTGGATCACAGATGCCATT 482
 421 CTCGCGCAACCGCTCAGTGGGCTGATCACTTAATCTCCGCTGGATCACAGATGCCATT 480

QY 483 GCTGTGGAAGCTGCTGCACATAATGTTCCGCGGTATTCTTGATGTCCTGACAGACA 542
 DB 481 GCTGTGGAAGCTGCTGCACATAATGTTCCGCGGTATTCTTGATGTCCTGACAGACA 540
 QY 543 CCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGGGCTGGGAGCATCTG 602
 DB 541 CCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGGGCTGGGAGCATCTG 600
 QY 603 GTCCATTGGGTACACCAAAATCGCGCTGTAGCGGCGCCATTAAATTTCTGTCGCGC 662
 DB 601 GTCCATTGGGTACACCAAAATCGCGCTGTAGCGGCGCCATTAAATTTCTGTCGCGC 660
 QY 663 CGTCTGCGTCTGCGTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCGATAGCG 722
 DB 661 CGTCTGCGTCTGCGTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCGATAGCG 720
 QY 723 GAACGGGAAGCGGACCTGAGTGCATGTCGGTTCACAAACCAACCATGCAATGCTGAAT 782
 DB 721 GAACGGGAAGCGGACCTGAGTGCATGTCGGTTCACAAACCAACCATGCAATGCTGAAT 778
 QY 783 GAGG 786
 DB 779 GAGG 782

RESULT 5

LOCUS

DEFINITION

BQ751220 806 bp mRNA linear EST 18-JUL-2002
 EST631783 DSC7 Colletotrichum trifolii cDNA clone pDSC73-64, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Colletotrichum trifolii
 Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes; incertae sedis; Phyllachorales; Phyllachoraceae;
 mitosporic Phyllachoraceae; Colletotrichum.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 806)
 Samac, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
 Cheung, F. and Fraser, C.M.
 ESTs from mycelia of Colletotrichum trifolii race 1
 Unpublished (2002)
 Other ESTs: EST631782

Contact: Deborah A. Samac

Department of Plant Pathology

University of Minnesota

495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA

Tel: 612 625 1243

Fax: 651 649 5058

Email: debbys@puccini.crl.umn.edu

TIGR sequence name: MTSAC64TV More information is available at:

www.medicago.org

Seq primer: (gtA AtA CgA CtC Act AtA ggg C).

FEATURES

source

Location/Qualifiers
 1..806
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDSC73-64"
 /tissue_type="mycelia"
 /dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSC7"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl1 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid

lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK-. Aliquots of the ligation were used to transform *E. coli* Disalpa which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match	22.9%	Score 755.4	DB 5	Length 806
Best Local Similarity	97.4%	Pred. No. 2.5e-212		
Matches 790	Conservative 0	Mismatches 16	Indels 5	Gaps 2
QY	3	GACACCATCAATGGTGCAGAAACCTTTCCGGGTATGCGATATGCGATAGCGCCCGGAAGAGAGT	62	
DB	1	GACATTTCCGATGGCGCAAAACCTTTTCGGGTATGCGATATGCGATAGCGCCCGGAAGAGAGT	60	
QY	63	CAATTCAGGGTGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCGGT	122	
DB	61	CAATTCAGGGTGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCGGT	120	
QY	123	GTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACAGCGCCAGCCACGTTTCTGCGAACAACG	182	
DB	121	GTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACAGCGCCAGCCACGTTTCTGCGAACAACG	180	
QY	183	CGGGAACAGTGAAGCGCGGATGCGCGAGCTGAATTAACATTCGCCAACCGCGTGGCACA	242	
DB	181	CGGGAACAGTGAAGCGCGGATGCGCGAGCTGAATTAACATTCGCCAACCGCGTGGCACA	240	
QY	243	CAACTGGCGGGCAACAGTCTGTTCTGATTTGGCGTTGCCACCTCCAGTCTGGGCCCTGCAC	302	
DB	241	CAACTGGCGGGCAACAGTCTGTTCTGATTTGGCGTTGCCACCTCCAGTCTGGGCCCTGCAC	300	
QY	303	CGCGCGTGCAGAAATTTGTCGGCGCGATTAATCTCGCGCGCATCAATCGGGTGCACAGCGTG	362	
DB	301	CGCGCGTGCAGAAATTTGTCGGCGCGATTAATCTCGCGCGCATCAATCGGGTGCACAGCGTG	360	
QY	363	GTGGTGTGCATGTGTAGAACGAAGCGGCGTGCAGCGCTGTAAAGCGGCGGTGCACAATCTT	422	
DB	361	GTGGTGTGCATGTGTAGAACGAAGCGGCGTGCAGCGCTGTAAAGCGGCGGTGCACAATCTT	420	
QY	423	CTCGCGCAACGCGTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCGAGGATGCCATT	482	
DB	421	CTCGCGCAACGCGTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCGAGGATGCCATT	480	
QY	483	GCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACCGAGACA	542	
DB	481	GCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACCGAGACA	540	
QY	543	CCCATCAACAGTATATTTCTCCATGAAGACGGTAGCGACTGGGGGTGGAGCATCTG	602	
DB	541	CCCATCAACAGTATATTTCTCCATGAAGACGGTAGCGACTGGGGGTGGAGCATCTG	600	
QY	603	GTCCGATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGGCG	662	
DB	601	GTCCGATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGGCG	660	
QY	663	CGTCTGGCTCTGGCTGGCTGCGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG	722	
DB	661	CGTCTGGCTCTGGCTGGCTGCGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG	720	
QY	723	GACCGGAAGCGCACTCGAGTGCCATGTCGGGTTTTCAACAAACCATGCAAAATCTCTCAAT	782	
DB	721	GACCGGAAGCGCACTCGAGTGCCATGTCGGGTTTTCAACAAACCATGCAAAATCTCTCAAT	777	
QY	783	GAGGGCATCGTTCCCACTGCGATGCTGGTTG	813	
DB	778	GAGG- -CATCGTTCCACTGCGATGCTGGTTG	806	

RESULT 6

CB863541/c

LOCUS	CB863541	689 bp	mrna	linear	EST	22-APR-2003
DEFINITION	HH04A08Y	HH Hordeum vulgare	cDNA clone	HH04A08	3-PRIME,	mrna
						sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE AUTHORS

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1965, Vol. 68, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1978, Vol. 1, No. 2, pp. 1-10
3. The Importance of Parental Involvement	Journal of Educational Psychology	1985, Vol. 77, No. 3, pp. 1-10
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1990, Vol. 93, No. 1, pp. 1-10
5. The Role of the School in the Community	Journal of Educational Research	1995, Vol. 98, No. 1, pp. 1-10
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	2000, Vol. 103, No. 1, pp. 1-10
7. The Role of the Teacher in the Classroom	Journal of Educational Research	2005, Vol. 108, No. 1, pp. 1-10
8. The Impact of Technology on Education	Journal of Educational Technology	2010, Vol. 3, No. 2, pp. 1-10
9. The Importance of Parental Involvement	Journal of Educational Psychology	2015, Vol. 87, No. 3, pp. 1-10
10. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	2020, Vol. 123, No. 1, pp. 1-10

FEATURES
SOURCE

CE863541
 CB863541.1 GI:30058100
 EST.
 Hordeum vulgare
 Hordeum vulgare
 Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.
 1 (bases 1 to 689)
 Varsheina,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
 Graner,A.
 Barley ESTs from coleoptile tissue
 Unpublished (2003)
 Contact: Stein Nils
 Molecular Markers Group, Department GenBank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3, 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5535
 Email: stein@ipk-gatersleben.de
 Insert Length: 689 Std Error: 0.00
 Plate: 4 row: A column: 8
 Seg primer: SP6.

FEATURES

```

1. .689
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555198"
/db_xref="taxon.4513"
/clone="HH04A08"
/tissue_type="coleoptile"
/dev_stage="coleoptile, 1 day old"
/lab_host="DH10B"
/clone_lib="HH"
/note="Vector: pSPORT; Site_1: SalI (5-end of cDNA);
Site_2: NotI (3-end of cDNA); Due to the cloning system
used blue/white selection for recombinants is not 100 %
reliable. Average insert size is 1.3 kb."

```

ORIGIN

Query Match 20.8%; Score 684.8; DB 6; Length 689;
Best Local Similarity 99.6%; Pred. No. 2.3e-191;
Matches 686: Conservative 0; Mismatches 3; Indels 0; Gaps 0

Qy	21	AAAACTTTTCGCGGTATGCGATGATGAGCGCCGCGAAGAGAGTCAATTCAGGGTGGTGAAT	80
Db	689	AAAACTTTTCGCGGTATGCGATGATGAGCGCCGCGAAGAGAGTCAATTCAGGGTGGTGAAT	630
Qy	81	GTGAACACAGTAAACGTTATACGATGTCGACAGTATGCCGGTGTCCTCTTATCGACACGGTT	140
Db	629	GTGAACACAGTAAACGTTATACGATGTCGACAGTATGCCGGTGTCCTCTTATCGACACGGTT	570
Qy	141	TCCCGCGTGGTGAAACAGCGCCAGCCACGCTTTCTCGAAAAACGCGGAAAAAGTGCGAAGCG	200
Db	569	TCCCGCGTGGTGAAACAGCGCCAGCCACGCTTTCTCGAAAAACGCGGAAAAAGTGCGAAGCG	510
Qy	201	GCATGGCGGAGCTGAATTATATCCCAACCGCGTGGGCACAACTTGGCGGGCAAAACAG	260
Db	509	GCATGGCGGAGCTGAATTATATCCCAACCGCGTGGGCACAACTTGGCGGGCAAAACAG	450
Qy	261	TGCTTGCTGATTGGGGTTGGCCACCTCCAGTCTGGCCCTGCACGCGCGTGCCTCAATTTGTC	320
Db	449	TGCTTGCTGATTGGGGTTGGCCACCTCCAGTCTGGCCCTGCACGCGCGTGCCTCAATTTGTC	390
Qy	321	CGCGCGATTAAATCTCGCGCCGATCAACTGGGTGTCACGCGTGGTGGTGTGCATGGTAGAA	380
Db	389	CGCGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTGCATGGTAGAA	330
Qy	381	CGAAGCGGCGTGCAGACCTGTAAACGGCGGGTGCACAACTCTTCTCGCGCCAACCGCGTCAGT	440
Db	329	CGAAGCGGCGTGCAGACCTGTAAACGGCGGGTGCACAACTCTTCTCGCGCCAACCGCGTCAGT	270

```

QY 441 GGGCTGATCAATTAATATCCGCTGGATGACACAGGATGCCATTGCTGTGTGAAGCTGCCTGC 500
Db |||||
QY 269 GGGCTGATCAATTAATATCCGCTGGATGACACAGGATGCCATTGCTGTGTGAAGCTGCCTGC 210
Db |||||
QY 501 ACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACAGACACCCATCAACAGTATTATT 560
Db |||||
QY 209 ACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACAGACACCCATCAACAGTATTATT 150
QY 561 TTCTCCCATGAAGACGGTACGGACTGCGCGTGGAGCATCTGGTCGCATTGGGTCACACAG 620
Db |||||
QY 149 TTCTCCCATGAAGACGGTACGGACTGCGCGTGGAGCATCTGGTCGCATTGGGTCACACAG 90
QY 621 CAATTCGGCGCTTTAGCGGGCCCAATTAAGTTCTGTCTCGGCGCGTCTCGCTGCGC 680
Db |||||
QY 89 CAAATCGCGCTTTAGCGGGCCCAATTAAGTTCTGTCTCGGCGCGTCTCGCTGCGCTGC 30
QY 681 TGGCATTAATATCTCACTCGCAATCAAT 709
Db |||||
29 TGGCATTAATATCTCACTCGCAATCAAT 1

RESULT 7
AL038548/c
LOCUS
DEFINITION
  DKFZps566E0346_r1 566 (synonym: hfkd2) Homo sapiens cDNA clone
  AL038548
  AL038548
  AL038548.1 GI:5407738
  EST.
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 712)
  Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
  Wiemann,S.
  EST (Ottenwaelder, et al.)
  Unpublished (1999)
  Contact: MIPS
  MIPS
  Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
  This is the 5' sequence of the clone insert
  Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
  Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
  sequenced by MediGenomix (Martinried/Germany) within the cDNA
  sequencing consortium of the German Genome Project. No s1 sequence
  available.
  This clone (DKFZps566E0346) is available at the RZPD in Berlin.
  Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
  Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
  Location/Qualifiers
    1..712
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="DKFZps566E0346"
      /tissue_type="kidney"
      /dev_stage="fetal"
      /lab_host="X1-2blue"
      /clone_lib="566 (synonym: hfkd2)"
      /note="vector: pAMP1; Site_1: NotI; Site_2: SalI"

FEATURES
  source
    1..712
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="DKFZps566E0346"
      /tissue_type="kidney"
      /dev_stage="fetal"
      /lab_host="X1-2blue"
      /clone_lib="566 (synonym: hfkd2)"
      /note="vector: pAMP1; Site_1: NotI; Site_2: SalI"

ORIGIN
  Query Match 20.5%; Score 676.4; DB 1; Length 712;
  Best Local Similarity 99.6%; Pred. No. 7.2e-189;
  Matches 699; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 33 GGTATGCGATGATAGCCCGGAGAGAGATCAATTCAGGTGGTGTGATGTGAACACAGTA 92
Db |||||
QY 712 GGTATGCGATGATAGCCCGGAGAGAGATCAATTCAGGTGGTGTGATGTGAACACAGTA 653
QY 93 ACGTTATACATGTCGAGAGTAGTCGCGGTCTCTTTATCAGACCGGTTCCCGCGTGTG 152
Db |||||

```

```

Db 652 ACGTTATACATGTCGAGAGTAGTCGCGGTCTCTTTATCAGACCGTTTCCCGCGTGTG 593
QY 153 AACAGCGCAGCCAGTTTCTGCGAAACCGCGG-AAAAAGTGAAGCGCGATGGCGGA 211
Db |||||
QY 592 AACAGCGCAGCCAGTTTCTGCGAAACCGCGGAAAAAGTGAAGCGCGATGGCGGA 533
QY 212 GCTCAATTAACATTTCCCAACCGCGTGGCACAACACTGCGCGGCAACAGTGTGCTGAT 271
Db |||||
QY 532 GCTCAATTAACATTTCCCAACCGCGTGGCACAACACTGCGCGGCAACAGTGTGCTGAT 473
QY 272 TGGCGTTGCCACCTCCAGTCTGCGCCCTGCAC- GCGCCGTGCAAAATTTGTCGCGCGATTA 330
Db |||||
QY 472 TGGCGTTGCCACCTCCAGTCTGCGCCCTGCACGCGCGCTGCAAAATTTGTCGCGCGATTA 413
QY 331 AATCTCGCGCGATCAACTGCGTGCAGCGTGGTGTGTCGATGGTGAAGACGAAGCGCG 390
Db |||||
QY 412 AATCTCGCGCGATCAACTGCGTGCAGCGTGGTGTGTCGATGGTGAAGACGAAGCGCG 353
QY 391 TCGAAGCCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACCGGTGAGTGGGCTGATCA 450
Db |||||
QY 352 TCGAAGCCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACCGGTGAGTGGGCTGATCA 293
QY 451 TTAACATATCCGCTGGATGACAGGATGCCATTGCTGTGAAGCTGCTGCACCTAAATGTT 510
Db |||||
QY 292 TTAACATATCCGCTGGATGACAGGATGCCATTGCTGTGAAGCTGCTGCACCTAAATGTT 233
QY 511 CGGCGTTATTTCTGATGTCCTGACGACACCCATCAACAGTATTATTTCTCCCATG 570
Db |||||
QY 232 CGGCGTTATTTCTGATGTCCTGACGACACCCATCAACAGTATTATTTCTCCCATG 173
QY 571 AAGACGCTAGCCGATCGGCGTGGAGCATCTGTCGCAATTTGGGTCAACGCAAAATCGCG 630
Db |||||
QY 172 AAGACGCTAGCCGATCGGCGTGGAGCATCTGTCGATTTGGGTCAACGCAAAATCGCG 113
QY 631 TGTAGCGGCGCCATTAAGTTCTGTCGCGCGTCTGCGTCTGGTGGTGGCATATAAT 690
Db |||||
QY 112 TGTAGCGGCGCCATTAAGTTCTGTCGCGCGTCTGCGTCTGGTGGTGGCATATAAT 53
QY 691 ATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAG 732
Db |||||
52 ATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAG 11

RESULT 8
AY080106
LOCUS
DEFINITION
  AY080106 Scripps Pier (La Jolla, CA) uncultured virus community
  uncultured marine virus genomic clone SI05lp36L, genomic survey
  sequence.
  AY080106
  AY080106.1 GI:24745302
  GSS.
  uncultured marine virus
  uncultured marine virus
  Viruses; environmental samples.
  1 (bases 1 to 710)
  Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
  Mead,D., Azam,P. and Rohwer,F.
  Genomic analysis of uncultured marine viral communities
  Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
  22294988
  12384570
  Contact: Rohwer F
  Biology Dept.
  San Diego State University
  5500 Campanile Dr, San Diego, CA 92102, USA
  Tel: 6195941336
  Fax: 619595676
  Email: forest@unstroke.sdsu.edu
  Class: shotgun.
  Location/Qualifiers
    1..710
      /organism="uncultured marine virus"

FEATURES
  source
    1..710
      /organism="uncultured marine virus"

```

/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone="SI05ip36L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus community"
/note="Marine viruses were isolated from 200 liters of surface seawater using a combination of differential filtration and density-dependent gradient centrifugation. Linker-amplified shotgun libraries were constructed by randomly shearing the total marine viral community DNA, end-repairing, ligating dsDNA linkers to the ends, and amplifying the fragments using Vent DNA polymerase. The resulting fragments were ligated into the pSMART vector and electroporated into MC12 cells (Lucigen; Middleton, WI)."

ORIGIN

Query Match 20.5%; Score 676; DB 8; Length 710;
Best Local Similarity 98.9%; Pred. No. 9.5e-189;
Matches 690; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 400 GTAAAGCGCGTGACAAATCTTCTCGGCAACGGCTCAGTGGCTGATCAATTAATCACTATC 459
DB 8 GTGCTGACGGGTGCACAACTCTTCTCGGCAACGGCTCAGTGGCTGATCAATTAATCACTATC 67
QY 460 CGCTGGATGACAGGATGCAATGCTGTGGAAAGCTGCTGCACTAATGTTCCGGCGTTAT 519
DB 68 CGCTGGATGACAGGATGCAATGCTGTGGAAAGCTGCTGCACTAATGTTCCGGCGTTAT 127
QY 520 TCTTTGATGTCCTGACAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTA 579
DB 128 TCTTTGATGTCCTGACAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTA 187
QY 580 CGCGACTGGCGTGGAGCATCTGTGCGATTTGGTCCACGAAATCGCGCTGTTAGCGG 639
DB 188 CGCGACTGGCGTGGAGCATCTGTGCGATTTGGTCCACGAAATCGCGCTGTTAGCGG 247
QY 640 GCCATTAAGTTCTGTCTCGGGCGCTGTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 699
DB 248 GCCATTAAGTTCTGTCTCGGGCGCTGTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 307
QY 700 GCAATCAATTCAGCCGATAGCGGAACGGGAGGCGACTGGAGTCCCATGTCGGGTTTC 759
DB 308 GCAATCAATTCAGCCGATAGCGGAACGGGAGGCGACTGGAGTCCCATGTCGGGTTTC 367
QY 760 AACAAACCATGCAATGCTGAATGAGGCGCATCGTTCCCACTCGCATGCTGTTGCAACG 819
DB 368 AACAAACCATGCAATGCTGAATGAGGCGCATCGTTCCCACTCGCATGCTGTTGCAACG 427
QY 820 ATCAGATGGCGTGGCGCAATGCGCGCATTTACCGAGTTCGGGCTGCGCGTGTGTCGG 879
DB 428 ATCAGATGGCGTGGCGCAATGCGCGCATTTACCGAGTTCGGGCTGCGCGTGTGTCGG 487
QY 880 ATATCTCGGTAGTGAGTACGAGTACCGAAGACAGCTCATGTTATATCCCGCGTTAA 939
DB 488 ATATCTCGGTAGTGAGTACGAGTACCGAAGACAGCTCATGTTATATCCCGCGTTAA 547
QY 940 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACGAGCTGGACGCTGCTGTCGAAC 999
DB 548 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACGAGCTGGACGCTGCTGTCGAAC 607
QY 1000 TCTCTCAGGGCGAGCGGTGAAGGCAATTCAGCTGTGTCGCGTCTCACTGGTGAAGAA 1059
DB 608 TCTCTCAGGGCGAGCGGTGAAGGCAATTCAGCTGTGTCGCGTCTCACTGGTGAAGAA-AA 666
QY 1060 AAACACCTGCGGCCCAATACGCAACCGCCTCTCCC 1097
DB 667 AAACACCTGCGGCCCAATACGCAACCGCCTCTCCC 704

RESULT 9
AL044483/c
LOCUS

660 bp mRNA linear EST 04-SEP-2003

DEFINITION DKFZp4341102_s1 434 (synonym: htes3) Homo sapiens cDNA clone
DKFZp4341102_3', mRNA sequence.
ACCESSION AL044483
VERSION AL044483.1 GI:5432701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Ansoorge, W., Benes, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and Wismann, S.
TITLE EST (Ansoorge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wismann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wismann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp4341102) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..660
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp4341102"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 19.6%; Score 647.4; DB 1; Length 660;
Best Local Similarity 99.7%; Pred. No. 2.9e-180;
Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 85 AACAGTAACGTTATACGATGTCGAGATATGCGGTGTCTCTTATCAGACCGTTCCC 144
DB 660 AACAGTAACGTTATACGATGTCGAGATATGCGGTGTCTCTTATCAGACCGTTCCC 601
QY 145 CGTGTGTAAACAGCCAGCCAGCCAGTTTCTGCGAAACCGGGAAAAAGTGGAGCGGCA 204
DB 600 CGTGTGTAAACAGCCAGCCAGCCAGTTTCTGCGAAACCGGGAAAAAGTGGAGCGGCA 541
QY 205 TGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACATGCGCGGCAACAGTGT 264
DB 540 TGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACATGCGCGGCAACAGTGT 481
QY 265 TGCTGATTGGGTTGCCACCTCCAGTCTGGCCCTGTCACGCGCGTTCGCAAAATGTCGCGG 324
DB 480 TGCTGATTGGGTTGCCACCTCCAGTCTGGCCCTGTCACGCGCGTTCGCAAAATGTCGCGG 421
QY 325 CGATTAAATCTCGGCGCCGATCAACTGGGTGTCAGGTGGTGTGTCGATGTTAGAACGAA 384
DB 420 CGATTAAATCTCGGCGCCGATCAACTGGGTGTCAGGTGGTGTGTCGATGTTAGAACGAA 361
QY 385 GCGGCGTCGAAGCGTGAAGCGGGTGCACATCTTCTCGCGCAACGCGTCAGTGGGC 444
DB 360 GCGGCGTCGAAGCGTGAAGCGGGTGCACATCTTCTCGCGCAACGCGTCAGTGGGC 301
QY 445 TGATCATTAATCTATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTA 504
DB 300 TGATCATTAATCTATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTA 241
QY 505 ATGTTCCGGCGTTATTTCTTGATGTCCTGACGACACCCATCAACAGTATTATTTTCT 564

```

Db      240 ATGTTCCGGCGTATTCTTGATGTTCTGACGACACCCATCAACAGTATTATTTCT 181
QY      565 CCATGAAGACGGTAGCGGACTGGCGGTGGAGCATCTGTCATTTGGTCAACAGCAAA 624
Db      180 CCATGAAGACGGTAGCGGACTGGCGGTGGAGCATCTGTCGCAATTGGGTCAACAGCAA 121
QY      625 TCGCGTGTAGCGGCGCCCAATTAAGTTCTGTCCTCGCGGTCTGCGTCTGGGTGCTGCTGCG 684
Db      120 TCGCGGTGTAGCGGCGCCCAATTAAGTTCTGTCCTCGG- GGTCTGCGTCTGGGTGCTGCG 62
QY      685 ATAAATATCTACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCCGACTGGAGTG 744
Db      61 ATAAATATCTACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCCGACTGGAGTG 2
QY      745 C 745
Db      1 C 1

RESULT 10
BJ063872/c
LOCUS   BJ063872
DEFINITION BJ063872 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL077k02 5', mRNA sequence.
ACCESSION BJ063872
VERSION   BJ063872.1 GI:17471074
KEYWORDS EST.
SOURCE   Xenopus laevis (African clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 641)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
TITLE     Expressed genes in X. laevis embryo
JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-i
          Center For Genetic Resource Information
          National Institute of Genetics
          111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshini@genes.nig.ac.jp
          The information of this clone is available through the following URL.
          http://xenopus.nibb.ac.jp.
FEATURES             source
     Location/Qualifiers
         1..641
             /organism="Xenopus laevis"
             /mol_type="mRNA"
             /db_xref="taxon:8355"
             /clone="XL077k02"
             /tissue_type="whole embryo"
             /dev_stage="stage 25"
             /clone_lib="NIBB Mochii normalized Xenopus tailbud library"

ORIGIN
Query Match      19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred.No. 2.3e-178;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30 CGCGGTATGCATGATAGCCCGGAGAGAGATCAATTCAGGTGGTGAATGTGAACCA 89
Db      641 CGCGGTATGCATGATAGCCCGGAGAGAGATCAATTCAGGTGGTGAATGTGAACCA 582
QY      90 GTAACGTTATACGATGTCGAGAGATAGCCGGTGTCTCTTATCAGACCGTTTCCCGGTG 149
Db      581 GTAACGTTATACGATGTCGAGAGATAGCCGGTGTCTCTTATCAGACCGTTTCCCGGTG 522
QY      150 GTGAACACCGCCAGCCACGTTTCTCGCAAAACGCGGAAAAAGTGAAGCGCGATGGCG 209

```

```

Db      521 GTGAACACCGCCAGCCACGTTTCTCGCAAAACGCGGAAAAAGTGAAGCGCGATGGCG 462
QY      210 GAGCTGAATTACATTTCCAAACCGGTGGCAACAACTGGCGGCAACAGTCGTTGCTG 269
Db      461 GAGCTGAATTACATTTCCAAACCGGTGGCAACAACTGGCGGCAACAGTCGTTGCTG 402
QY      270 ATTGGCGTTGCACTTCAAGTCTGCGCCTGCAACGCGCGTTCGCAAAATTGTCGCGCGGAT 329
Db      401 ATTGGCGTTGCACTTCAAGTCTGCGCCTGCAACGCGCGTTCGCAAAATTGTCGCGCGGAT 342
QY      330 AAATCTCGCGCGGATCAACTGGGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 389
Db      341 AAATCTCGCGCGGATCAACTGGGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 282
QY      390 GTGAAGCCTGTAAAGCGCGGTGCACAAATCTTCTCGCGCAACGCGTCAGTCAGTCAGTCAGTC 449
Db      281 GTGAAGCCTGTAAAGCGCGGTGCACAAATCTTCTCGCGCAACGCGTCAGTCAGTCAGTCAGTC 222
QY      450 ATTAACCTATCGCTGGATGACAGGATGCCATTCGTTGTGGAAAGTCCTGCATTAATGTT 509
Db      221 ATTAACCTATCGCTGGATGACAGGATGCCATTCGTTGTGGAAAGTCCTGCATTAATGTT 162
QY      510 CCGCGGTATTCTTCTGATGTCCTGACCAACACCATCAACAGTATTATTTCTCCCAT 569
Db      161 CCGCGGTATTCTTCTGATGTCCTGACCAACACCATCAACAGTATTATTTCTCCCAT 102
QY      570 GAAGACGGTAGCGGACTGGCGGTGGAGCATCTGTCGCAATTCGCGTACCAGCAAAATCGCG 629
Db      101 GAAGACGGTAGCGGACTGGCGGTGGAGCATCTGTCGCAATTCGCGTACCAGCAAAATCGCG 42
QY      630 CTGTTAGCGGCGCCATTAAGTTCTGTCGCGCGCTCTCGG 670
Db      41 CTGTTAGCGGCGCCATTAAGTTCTGTCGCGCGCTCTCGG 1

RESULT 11
CB863814/c
LOCUS   CB863814
DEFINITION HH07B20y HH Hordeum vulgare cDNA clone HH07B20 3-PRIME, mRNA sequence.
ACCESSION CB863814
VERSION   CB863814.1 GI:30058373
KEYWORDS EST.
SOURCE   Hordeum vulgare
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 624)
AUTHORS Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and Graner,A.
TITLE     Barley ESTs from coleoptile tissue
JOURNAL   Unpublished (2003)
COMMENT   Contact: Stein Nils
          Molecular Markers Group, Department Genbank
          Institute of Plant Genetics and Crop Plant Research (IPK)
          Corrensstr. 3 06466, Gatersleben, Germany
          Tel: 039482-5522
          Fax: 039482-5595
          Email: stein@ipk-gatersleben.de
          Insert Length: 624 Std Error: 0.00
          Plate: 7 row: B column: 20
          Seq primer: SP6.
          Location/Qualifiers
             1..624
                 /organism="Hordeum vulgare"
                 /mol_type="mRNA"
                 /cultivar="Sloop"
                 /db_xref="GABI:555471"
                 /db_xref="taxon:4513"
                 /clone="HH07B20"
                 /tissue_type="coleoptile"

```


/dev_stage="coleoptile, 1 day old"
 /lab_host="DH10B"
 /clone_lib="HH"
 /note="Vector: pSPORT; Site 1: SalI (5-end of cDNA); Site 2: NotI (3-end of cDNA); Due to the cloning system used blue/white selection for recombinants is not 100 % reliable. Average insert size is 1.3 Kb."

ORIGIN
 Query Match 18.7%; Score 615.8; DB 6; Length 624;
 Best Local Similarity 99.7%; Pred. No. 7e-171;
 Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GACACATCGAATGGTGCAGAAACCTTTCCGGGTATGCGATGATAGCCCGGAGAGAGT 62
 Db |
 619 GACACATCGAATGGCGCAAAACCTTTCCGGGTATGCGATGATAGCCCGGAGAGAGT 560
 Qy 63 CAATTCAGGTGGTGAATGTGAACACAGTAACTGATACAGTATGCGAGAGTATCCGGT 122
 Db |
 559 CAATTCAGGTGGTGAATGTGAACACAGTAACTGATACAGTATGCGAGAGTATCCGGT 500
 Qy 123 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAACG 182
 Db |
 499 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAACG 440
 Qy 183 CGGAAAAATGGAAGCGCGATGGCGAGCTGAATTTACATTTCCCAACCCGCTGGCAAA 242
 Db |
 439 CGGAAAAATGGAAGCGCGATGGCGAGCTGAATTTACATTTCCCAACCCGCTGGCAAA 380
 Qy 243 CAATCGCGGCAACAGTCTGTTCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGCAC 302
 Db |
 379 CAATCGCGGCAACAGTCTGTTCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGCAC 320
 Qy 303 GCGCGTGCAGAAATTTGCGCGGCAATTAATTTCTCGCGCGATCAACTGGGTGCCAGCGTG 362
 Db |
 319 GCGCGTGCAGAAATTTGCGCGGCAATTAATTTCTCGCGCGATCAACTGGGTGCCAGCGTG 260
 Qy 363 GTGGTGTGATGTGAGAACGCGGCTCGAAGCCTGTAAAGCGCGGTGCACAAATCTT 422
 Db |
 259 GTGGTGTGATGTGAGAACGCGGCTCGAAGCCTGTAAAGCGCGGTGCACAAATCTT 200
 Qy 423 CTCGCGCAACGCTCAGTGGGTGATCAATTAATCTATCGGTGGATGACAGGATGCCATT 482
 Db |
 199 CTCGCGCAACGCTCAGTGGGTGATCAATTAATCTATCGGTGGATGACAGGATGCCATT 140
 Qy 483 GCTGTGAAGCTGCTCAGCTAATTTCCGCGTTATTTCTTGATGCTCTGACACACA 542
 Db |
 139 GCTGTGAAGCTGCTCAGCTAATTTCCGCGTTATTTCTTGATGCTCTGACACACA 80
 Qy 543 CCCATCAACAGTATTTTCTCCATGAAGACGCTGACGCGATCGGCGGTGGAGCATCTG 602
 Db |
 79 CCCATCAACAGTATTTTCTCCATGAAGACGCTGACGCGATCGGCGGTGGAGCATCTG 20
 Qy 603 GTCGATGGGTACACAGC 621
 Db |
 19 GTCGATGGGTACACAGC 1

RESULT 12
 AL045353/c
 LOCUS
 DEFINITION 794 bp mRNA linear EST 04-SEP-2003
 DKFP434B075.s1.434 (synonym: htes3) Homo sapiens cDNA clone
 DKFP434B075.3', mRNA sequence.
 AL045353
 AL045353.1 GI:5433506
 EST.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 794)
 Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and
 Wiemann,S.

FEATURES
 source
 1..794
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFP434B075"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI"

ORIGIN
 Query Match 18.5%; Score 612; DB 1; Length 794;
 Best Local Similarity 99.7%; Pred. No. 1e-169;
 Matches 634; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 569 TGAAGACGGTACCGGACTGGCGTGGAGCATCTGTGCGCATTTGGGTTCACAGCAAAATCGC 628
 Db |
 794 TGAAGACGGTACCGGACTGGCGTGGAG-ATCTGTGCGCATTTGGGTTCACAGCAAAATCGC 736
 Qy 629 GCTGTTAGCGGGCCCAATTAAGTTCTGCTCGCGCGCTCTGCGTCTGCGTGGCGCATAA 688
 Db |
 735 GCTGTTAGCGGGCCCAATTAAGTTCTGCTCGCGCGCTCTG-GTCTGGTGGCTGGCGATAA 677
 Qy 689 ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGCGACTTGGAGTGCAT 748
 Db |
 676 ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGCGACTTGGAGTGCAT 617
 Qy 749 GTCGGGTTTCAACAAACCATGCAATGCTGAATGAGGGCATGCTTCCCATCGCATGCT 808
 Db |
 616 GTCGGGTTTCAACAAACCATGCAATGCTGAATGAGGGCATGCTTCCCATCGCATGCT 557
 Qy 809 GGTTCGCAACGATCAGATGGCGTGGCGGCAATGCGCGCCATTTACCGAGTCCCGGGCTGG 868
 Db |
 556 GGTTCGCAACGATCAGATGGCGTGGCGGCAATGCGCGCCATTTACCGAGTCCCGGGCTGG 497
 Qy 869 CGTTGCTGGCGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTTATAT 928
 Db |
 496 CGTTGCTGGCGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTTATAT 437
 Qy 929 CCGCGGTTTAAACCAACCATCAAAACAGGATTTTCGCCCTGCTGGGGCAAAACCGCTGACCG 988
 Db |
 436 CCGCGGTTTAAACCAACCATCAAAACAGGATTTTCGCCCTGCTGGGGCAAAACCGCTGACCG 377
 Qy 989 CTTGCTGCAACTCTCTCAGGCGGCGGTGAAGGGCAATCAGCTGTTGCCCTCTCTCACT 1048
 Db |
 376 CTTGCTGCAACTCTCTCAGGCGGCGGTTGAAGGGCAATCAGCTGTTGCCCTCTCTCACT 317
 Qy 1049 GGTGAAGAAAAAACCAACCTGCGGCCCAATACGCAAAACCGCTCTCTCCCGCGCGTTGGC 1108
 Db |
 316 GGTGAAGAAAAAACCAACCTGCGGCCCAATACGCAAAACCGCTCTCTCCCGCGCGTTGGC 257
 Qy 1109 CGATTCAATTAATGAGTGGCAGCAGAGGTTTCCCGACTGGAAAGCGGGCAGGTGAGCGCA 1168
 Db |
 256 CGATTCAATTAATGAGTGGCAGCAGAGGTTTCCCGACTGGAAAGCGGGCAGGTGAGCGCA 197
 Qy 1169 ACGCAATTAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1204
 Db |
 196 ACGCAATTAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 161

RESULT 13
 CL655575/c
 LOCUS
 DEFINITION
 CL655575 752 bp DNA linear GSS 09-JUL-2004
 PRI0124a.C01 - PRI0124a.B21 (752) Mixed stage fosmid library of P.
 pacificus var. California Pristionchus pacificus genomic, genomic
 survey sequence.
 CL655575
 GSS
 CL655575.1 GI:50134897
 Pristionchus pacificus
 Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 1 (bases 1 to 752)
 Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
 AppADB: an AcedB database for the nematode satellite organism
 Pristionchus pacificus
 Nucleic Acids Res. 32 (1), D421-D422 (2004)
 Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.
 Seq primer: 77
 Class: fosmid ends.
 Location/Qualifiers
 1..752
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pBpifos-5 Fosmid vector"

Query Match 18.3%; Score 603; DB 9; Length 752;
 Best Local Similarity 98.9%; Pred. No. 4.7e-167;
 Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
 QY 1996 TTACCTGGCCGCTGATTGCTGTGAC -GGGGGTTATGCGTTCAAGTATGAAACGGCAA 2054
 DB 752 TTACCTGGCCGCTGATTGCTGTGACGGGGGTTATGCGTTCAAGTATGAAACGGCAA 693
 QY 2055 GTACGACATTAAAGACGTGGCGTGGATAACGCTGGCGGAAAGCGGGTCTGACCTTCCT 2114
 DB 692 GTACGACATTAAAGACGTGGCGTGGATAACGCTGGCGGAAAGCGGGTCTGACCTTCCT 633
 QY 2115 GTTGACCTGATTTAAACAAACACATGATGACACACCGATTACTCCATCGAGAAC 2174
 DB 632 GTTGACCTGATTTAAACAAACACATGATGACACACCGATTACTCCATCGAGAAC 573
 QY 2175 TGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGGGCATGGTCCCAACAT 2234
 DB 572 TGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGGGCATGGTCCCAACAT 513
 QY 2235 CGACACCGCAAGTGAATTTATGGTTAAGCGGTACTGCCGACCTTCAAGGGTCAACCATC 2294
 DB 512 CGACACCGCAAGTGAATTTATGGTTAAGCGGTACTGCCGACCTTCAAGGGTCAACCATC 453
 QY 2295 CAACACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGGCCGAGTCCGCAACAAAGCT 2354
 DB 452 CAACACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGGCCGAGTCCGCAACAAAGCT 393
 QY 2355 GCGAAAGAGTTCCTCGAAACCTATCTGTGATGAAGGTCTGGAAGCGGTTAA-TA 2413
 DB 392 GCGAAAGAGTTCCTCGAAACCTATCTGTGATGAAGGTCTGGAAGCGGTTAATTA 333

QY 2414 AAGACAAACCGCTGGTGGTCCGTAGCGTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATC 2473
 DB 332 AGGCCAACCCCGTTGGTGGTCCGTAGCGTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATC 273
 QY 2474 CACGTATTGGCGCCACCATGAAACCCCGAGAAAGTGAAATCATCGCGAACATCCCGC 2533
 DB 272 CACGTATTGGCGCCACCATGAAACCCCGAGAAAGTGAAATCATCGCGAACATCCCGC 213
 QY 2534 AGATGTCGGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCCCGCGCGGTGTC 2593
 DB 212 AGATGTCGGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCCCGCGCGGTGTC 153
 QY 2594 AGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT 2628
 DB 152 AGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT 118
 RESULT 14
 AY080094/c
 LOCUS
 DEFINITION
 AY080094 608 bp DNA linear GSS 06-NOV-2002
 uncultured marine virus genomic clone SIO51p3D4L, genomic survey
 sequence.
 AY080094
 ACCESSION
 AY080094.1 GI:24745276
 GSS.
 KEYWORDS
 uncultured marine virus
 SOURCE
 uncultured marine virus
 ORGANISM
 Viruses; environmental samples.
 REFERENCE
 1 (bases 1 to 608)
 AUTHORS
 Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
 Mead,D., Azam,F. and Rohwer,F.
 TITLE
 Genomic analysis of uncultured marine viral communities
 JOURNAL
 Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
 MEDLINE
 22294988
 PUBMED
 12384570
 COMMENT
 Contact: Rohwer F
 Biology Dept.
 San Diego State University
 5500 Campanile Dr, San Diego, CA 92102, USA
 Tel: 6195941336
 Fax: 619595676
 Email: forest@sunstroke.sdsu.edu
 Class: shotgun.
 Location/Qualifiers
 1..608
 /organism="uncultured marine virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:186617"
 /clone="SIO51p3D4L"
 /clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
 community"
 /note="Marine viruses were isolated from 200 liters of
 surface seawater using a combination of differential
 filtration and density-dependent gradient centrifugation.
 Linker-amplified shotgun libraries were constructed by
 randomly shearing the total marine viral community DNA,
 end-repairing, ligating dsDNA linkers to the ends, and
 amplifying the fragments using Vent DNA polymerase. The
 resulting fragments were ligated into the pSMART vector
 and electroporated into MC12 cells (Lucigen; Middleton,
 WI)."

ORIGIN
 Query Match 17.9%; Score 589.8; DB 8; Length 608;
 Best Local Similarity 99.7%; Pred. No. 3.7e-163;
 Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 287 CAGTCTGGCCCTGACGGCCGCTGCAAAATTCGCGGCGATTAAATCTCGCGCGATCA 346
 DB 608 CAGTCTGGCCCTGACGGCCGCTGCAAAATTCGCGGCGATTAAATCTCGCGCGATCA 549

Qy	347	ACTGGGTGCCAGCTGGTGGTGTGATGTAGAACGAAGCGCGTCGAAGCCTGTAAAGC	406
Db	548	ACTGGGTGCCAGCGTGGTGGTGTGATGTAGAACGAAGCGCGTCGAAGCCTGTAAAGC	489
Qy	407	GGCGGTGCACAATCTTCTCGCGCAACGGTCAGTGGGCTGATCAATACTATCCGCTGA	466
Db	488	GGCGGTGCACAATCTTCTCGCGCAACGGTCAGTGGGCTGATCAATACTATCCGCTGA	429
Qy	467	TGACCAGGATGCCATTGCTGTGGAAAGCTGCCCTGCACTAATGTTCCGGCGTATATTTCTTCA	526
Db	428	TGACCAGGATGCCATTGCTGTGGAAAGCTGCCCTGCACTAATGTTCCAGCGTATATTTCTTGA	369
Qy	527	TGTTCTCTGACCAAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACT	586
Db	368	TGTTCTCTGACCAAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACT	309
Qy	587	GGGCGTGGAGCATCTGGTTCGCATTGGGTGCACAGCAAAATCGCGCTGTATAGCGGGCCCATTT	646
Db	308	GGGCGTGGAGCATCTGGTTCGCATTGGGTGCACAGCAAAATCGCGCTGTATAGCGGGCCCATTT	249
Qy	647	AAGTTCTGCTCTCGCGCGCTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCA	706
Db	248	AAGTTCTGCTCTCGCGCGCTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCA	189
Qy	707	AATTACGCCGATAGCGGAAACGGGAAGCGACATGGAGTGCATGTCGGGTTTTTCAACAAAC	766
Db	188	AATTACGCCGATAGCGGAAACGGGAAGCGACATGGAGTGCATGTCGGGTTTTTCAACAAAC	129
Qy	767	CATGCAAAATGCTGAATAGGGGCATGTTTCCCACTCGCATGTGTTTGGCCAAACGATCAGAT	826
Db	128	CATGCAAAATGCTGAATAGGGGCATGTTTCCCACTCGCATGTGTTTGGCCAAACGATCAGAT	69
Qy	827	GGCGCTGGGGCGAATGCGGGCCATTACGAGTCCGGGCTCGCGTTTGGTGGCGG	879
Db	68	GGCGCTGGGGCGAATGCGGGCCATTACGAGTCCGGGCTCGCGTTTGGTGGCGG	16

RESULT 15	AL037742	589 bp	mRNA	linear	EST 04-SEP-2003			
AL037742/c	DKFP2564C157.r1.564	(synonym: hfbr2)	Homo sapiens	cdna clone				
LOCUS	DKFP2564C157.5	mRNA sequence.						
DEFINITION	AL037742							
ACCESSION	AL037742.1	GI:5407088						
VERSION	EST.							
KEYWORDS	Homo sapiens							
SOURCE	Homo sapiens (human)							
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE	1 (bases 1 to 589)							
AUTHORS	Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and Wiemann,S.							
TITLE	EST (Bloecker, et al.)							
JOURNAL	Unpublished (1999)							
COMMENT	Contact: MIPS							
	MTS							

ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp564C157) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

```
FEATURES
source
1. 589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cclone="DKFZP564C157"
Location/Qualifiers
BERLIN-CHARLOTTENBURG, GENCODE;
```

```
/tissue_type="brain"  
/dev_stage="fetal"  
/lab_host="X1-2blue"  
/clone_lib="564 (synonym: hfbr2)"  
note="Vector: pAMP1; Site 1: NotI; Site 2: SalI"
```

ORIGIN

Query Match	17.8%;	Score 586.4;	DB 1;	Length 589;
Best Local Similarity	99.7%;	Pred. No. 3.7e-162;		
Matches 587;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	517	TATTTCTTGATGTC	CTGACAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACG	576
Db	589	TATTTCTTGATGTC	CTGACAGACACCCATCAACAGTATTATTTTCTCCCTTGAAGACG	530
Qy	577	GTACGGCACTGGGCGTGGAGCATCTGTCGCAATTGGTCCACGACAAATCGCGCTGTTAG	636	
Db	529	GTACGGCACTGGGCGTGGAGCATCTGTCGCAATTGGTCCACGACAAATCGCGCTGTTAG	470	
Qy	637	CGGGCCCAATTAAGTTCTGTCTCGCGCGCTCTGCGTCTGGCTGGCATTAATAATCTCA	696	
Db	469	CGGGCCCAATTAAGTTCTGTCTCGCGCGCTCTGCGTCTGGCTGGCATTAATAATCTCA	410	
Qy	697	CTCGCAATCAAAATTCACGCCATACGGGAAACGGGAAGCGCACTGGAGTGCCATGTCGCGTT	756	
Db	409	CTCGCAATCAAAATTCACGCCATACGGGAAACGGGAAGCGCACTGGAGTGCCATGTCGCGTT	350	
Qy	757	TTCACAAAACATGCAAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGTTGCCA	816	
Db	349	TTCACAAACCATGCAAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGTTGCCA	290	
Qy	817	ACGATCAGATGGCGCTGGGCGCAATCGCGCCATTACCGAGTCCGGGCTCGCGCTGGTG	876	
Db	289	ACGATCAGATGGCGCTGGGCGCAATCGCGCCATTACCGAGTCCGGGCTGGCGTTGGTG	230	
Qy	877	CGGATATCTCGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT	936	
Db	229	CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT	170	
Qy	937	TAACCAACATCAACAGAGTTTTTCGCCCTGCTGGGGCAACACGAGCGTGACCGCTTGTCG	996	
Db	169	TAACCAACATCAACAGAGTTTTTCGCCCTGCTGGGGCAACACGAGCGTGACCGCTTGTCG	110	
Qy	997	AACCTCTCTCAGGGCCAGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAA	1056	
Db	109	AACCTCTCTCAGGGCCAGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAA	50	
Qy	1057	GA AAAACCAACCTGGCGCCCAATAACGAAAACCGCCTCTCCCGCGCGTT	1105	
Db	49	GA AAAACCAACCTGGCGCCCAATAACGAAAACCGCCTCTCCCGCGCGTT	1	

Search completed: August 3, 2005, 02:02:46
Job time : 6748.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds
(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgacacatcgatgtgc.....acgacgttcggactaagct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3300	100.0	3300	10	US-09-765-555-14
2	3300	100.0	3300	10	US-09-765-555-15
3	3226.4	97.8	3300	10	US-09-765-555-18
4	3215.2	97.4	3300	10	US-09-765-555-17
5	3207.2	97.2	3300	10	US-09-765-555-16
6	2705.8	82.0	7373	18	US-09-765-384A-1
7	2705.8	82.0	8101	17	US-10-149-472-5

8	2701	81.8	6648	19	US-10-343-859-8	Sequence 8, Appli
9	2698	81.8	9191	13	US-10-343-859-9	Sequence 9, Appli
10	2690.4	81.5	7112	18	US-10-263-153-40	Sequence 40, Appl
11	2690.4	81.5	7259	18	US-10-263-153-30	Sequence 30, Appl
12	2690.4	81.5	7322	18	US-10-263-153-35	Sequence 35, Appl
13	2690.4	81.5	7352	18	US-10-263-153-25	Sequence 25, Appl
14	2690.4	81.5	7370	18	US-10-263-153-20	Sequence 20, Appl
15	2690.4	81.5	7370	18	US-10-263-153-61	Sequence 61, Appl
16	2690.4	81.5	7370	18	US-10-263-153-66	Sequence 66, Appl
17	2690.4	81.5	7370	18	US-10-263-153-71	Sequence 71, Appl
18	2690.4	81.5	7403	18	US-10-263-153-15	Sequence 15, Appl
19	2690.4	81.5	7442	18	US-10-263-153-10	Sequence 10, Appl
20	2690.4	81.5	7478	18	US-10-263-153-3	Sequence 3, Appli
21	2636.8	79.9	6806	10	US-09-908-943A-194	Sequence 194, App
22	2636.8	79.9	6806	20	US-10-801-487-194	Sequence 194, App
23	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
24	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
25	2636.8	79.9	6806	20	US-10-801-486-194	Sequence 194, App
26	2636.8	79.9	6806	21	US-10-801-493-194	Sequence 194, App
27	2605.4	79.0	7553	18	US-10-263-153-7	Sequence 7, Appli
28	1609.4	48.8	5558	16	US-10-241-596-137	Sequence 137, App
29	1474	44.7	4700	17	US-10-313-963A-54	Sequence 54, Appl
30	1462.4	44.3	4921	21	US-10-495-491-2	Sequence 2, Appli
31	1462.4	44.3	4935	21	US-10-495-491-1	Sequence 1, Appli
32	1462.4	44.3	4945	21	US-10-495-491-3	Sequence 3, Appli
33	1462.4	44.3	4951	21	US-10-495-491-4	Sequence 4, Appli
34	1417.6	43.0	1511	10	US-09-987-763-30	Sequence 30, Appl
35	1417.6	43.0	5926	13	US-10-024-809-3	Sequence 3, Appli
36	1238.4	37.5	5024	17	US-10-313-963A-52	Sequence 52, Appl
37	1238.4	37.5	5448	21	US-10-671-995A-2	Sequence 2, Appli
38	1238.4	37.5	5454	21	US-10-671-995A-1	Sequence 1, Appli
39	1237.2	37.5	1922	10	US-09-995-988A-32	Sequence 32, Appl
40	1237.2	37.5	1922	21	US-10-420-034A-32	Sequence 32, Appl
41	1237.2	37.5	6270	21	US-10-855-897-2	Sequence 2, Appli
42	1204	36.5	4557	17	US-10-457-372-5	GENERAL INFORMA
43	1204	36.5	6664	18	US-10-270-176-1	Sequence 1, Appli
44	1204	36.5	6668	18	US-10-270-176-2	Sequence 2, Appli
45	1204	36.5	6742	18	US-10-270-176-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-765-555-14
; Sequence 14, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m1 and zinc finger
; OTHER INFORMATION: Protein ZFPm1
US-09-765-555-14

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGCGCATGATAGCGCCGGGAAGAGA 60
Db 1 CGGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGCGCATGATAGCGCCGGGAAGAGA 60
QY 61 GTC AATTCAGGGTGGTGAATGTGA AACCAAGTAA CTTTATACGATGTCCGAGAGTATGCCG 120
Db 61 GTC AATTCAGGGTGGTGAATGTGA AACCAAGTAA CTTTATACGATGTCCGAGAGTATGCCG 120
QY 121 GTGTCCTCTTATCAGACCGTTTCCCGCGTGGTGAA CCGAGCGCAGCCAGCGTTTCTCGGAAA 180
Db 121 GTGTCCTCTTATCAGACCGTTTCCCGCGTGGTGAA CCGAGCGCAGCCAGCGTTTCTCGGAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTA CATTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTA CATTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCMAACAGTCTGTTGCTGATTTGGGTTTGCACCTTCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCMAACAGTCTGTTGCTGATTTGGCGTTTGCACCTTCAGTCTGGCCCTGC 300
QY 301 ACGGCGCGTCCGAAATTTGTCGGCGGATTA AATCTCGCGCGATCAACTGGGTGGCAGCG 360
Db 301 ACGGCGCGTCCGAAATTTGTCGGCGGATTA AATCTCGCGCGATCAACTGGGTGGCAGCG 360
QY 361 TGGTGGTGTGATGGTAGAACGAAGCGCGCTG AAGCGCTGAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTGATGGTAGAACGAAGCGCGCTG AAGCGCTGAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGGTCTAGTGGGCTGATCA TTAATCTATCCCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACGGTCTAGTGGGCTGATCA TTAATCTATCCCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCACTTAATGTTCCGGCGTTATTTCTTGATGCTCTCAGCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTTAATGTTCCGGCGTTATTTCTTGATGCTCTCAGCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTCCGATTTGGGTACCAAGCAATCGCGCTGTTAGCGGGCCCATTAAGTCTCTCTCGG 660
Db 601 TGGTCCGATTTGGGTACCAAGCAATCGCGCTGTTAGCGGGCCCATTAAGTCTCTCTCTCGG 660
QY 661 CGCGTCTCGCTCGCTGGCATAAATATCTCACTCGCAATCA AATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCA AATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGCGACTGGAGTGCCATGTC CGGTTTTCACAAACCATGCAATGCTGA 780
Db 721 CGGAAACGGGAAGCGACTGGAGTGCCATGTC CGGTTTTCACAAACCATGCAATGCTGA 780
QY 781 ATGAGGSCATCGTTCACACTCGATGCTGGTGC AACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGSCATCGTTCACACTCGATGCTGGTGC AACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGC CGCATTTACCGAGTCCGGGCTGGCGGTGGTGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGC CGCATTTACCGAGTCCGGGCTGGCGGTGGTGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACGGAAGACAGTCTATGTTATATCCGCGCGTTAA CCAATCAACAGGATTTTC 960
Db 901 ACGATACGGAAGACAGTCTATGTTATATCCGCGCGTTAA CCAATCAACAGGATTTTC 960
QY 961 GCCTGTGGGCGAAACAGCGTGGACCGCTTGCTGCAACTCTC CAGGCGCAGCGGTGA 1020
Db 961 GCCTGTGGGCGAAACAGCGTGGACCGCTTGCTGCAACTCTC CAGGCGCAGCGGTGA 1020
QY 1021 AGGCAATACAGCTGTTCCCGTCTCACTTGTTGA AAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATACAGCTGTTCCCGTCTCACTTGTTGA AAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCGCTCTCCCGCGCGTTGGCGGATTCATTAATG CAGCTGGCAGCAGAGTTT 1140

Db 1081 CGCAACCGCGCTCTCCCGCGCGTTGGCGGATTCATTAATG CAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATG AGTTAGTCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATG AGTTAGTCTCACTCATTTAG 1200
QY 1201 GCACAATCTCATGTTTTGACAGCTTATCATCGACTGCA CGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTTGACAGCTTATCATCGACTGCA CGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAAGCCATCGGAAAGCTGTGATGCGTGTGACAG GTCGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGCAAGCCATCGGAAAGCTGTGATGCGTGTGACAG GTCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGGGCGACTCCCGTTCTGGATAATATGTTTT TTTGCGCGCAGCATATAACGGTT 1380
Db 1321 TGTGCTCAAGGGCGACTCCCGTTCTGGATAATATGTTTT TTTGCGCGCAGCATATAACGGTT 1380
QY 1381 CTGCAATAATCTGAAATGAGCTGTTGACAAATTAATCAT CGGCTCGTATAATGTTGGA 1440
Db 1381 CTGCAATAATCTGAAATGAGCTGTTGACAAATTAATCAT CGGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTCA CAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTCA CAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
QY 1501 GCACCTTCA CAAACAGGACCATAGATTAATGAAATCTGAAGAGTAAATCTGGTAATCTGG 1560
Db 1501 GCACCTTCA CAAACAGGACCATAGATTAATGAAATCTGAAGAGTAAATCTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTTC GGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTTC GGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAA ACTGGAAGAGAAATTTCCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAA ACTGGAAGAGAAATTTCCACACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCA CACGACCGTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCA CACGACCGTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTCGCCTGTTGGCTGAAATCACCCCGA CAAAGCGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTCGCCTGTTGGCTGAAATCACCCCGA CAAAGCGTTCCAGGACAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTACACGCGCAAGCTGA TTGCTTACC CGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTACACGCGCAAGCTGA TTGCTTACC CGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCC GAACCGCCGCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCC GAACCGCCGCAAAAACCTGGGAA 1920
QY 1921 GAGATCCGCGCTGGATAAAGAACTGAAAGCAAAGGTAA GAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGGATAAAGAACTGAAAGCAAAGGTAA GAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGCGCTGATTTGCTGTA CCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGCGCTGATTTGCTGTA CCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGGCAAGTACGACATTAAGAGCTGGCGTGGAT AAACGCTGGCGGGAAGCG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTAAGAGCTGGCGTGGAT AAACGCTGGCGGGAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAAA CAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAAA CAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGGAGAGCTGCTCTTTTAAAGGCGAAACAGCGATG ACCATCAACGCGCGGTG 2220

Db 2161 TCCATCGCAGAGCTGCCCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCCCGGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTATGGTGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTATGGTGTAAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCAACCATCAACACCGTTCGTTGGCGTGTGAGGCGAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTCAACCATCAACACCGTTCGTTGGCGTGTGAGGCGAGGTATTAAACGCCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGAAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGAAAG 2460
Qy 2461 TTGGCGAAAGATCCACGTATTGCGCGCACCATGGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTGCGCGCACCATGGAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACAATCCGACAGTGTCCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACAATCCGACAGTGTCCGCTTTCTGGTATGCCGTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGCTGACAGTGTGATGAAGCCCTGAAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGACAGTGTGATGAAGCCCTGAAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACACAAACAAATAACAACTCCGGGATCGAGGAGAGGATTTTCAGAAATTC 2700
Db 2641 AACACACAAACAAATAACAACTCCGGGATCGAGGAGAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCACACAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATCTGGCGAGAGGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATCTGGCGAGAGGCCATACAAATGTCCAGATGTGGC 2940
Qy 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCGGCACCAACGTAATCTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCGGCACCAACGTAATCTCACACCGGGGAGAG 3000
Qy 3001 CCCTATGTGTTCGGAATGTGGTAAGTCTTTCAGCCGAGCGATAAATCTGGTGGCCAC 3060
Db 3001 CCCTATGTGTTCGGAATGTGGTAAGTCTTTCAGCCGAGCGATAAATCTGGTGGCCAC 3060
Qy 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
Qy 3121 CAGCGCGGCACCTGGCCAGGCATCAACGCACTCTACTCTGGCGAGAGGCCATACAAATGT 3180
Db 3121 CAGCGCGGCACCTGGCCAGGCATCAACGCACTCTACTCTGGCGAGAGGCCATACAAATGT 3180
Qy 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGCAATCTCGTCCGCGACCAACGTAATCTAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGCAATCTCGTCCGCGACCAACGTAATCTAC 3240
Qy 3241 ACCGGTAAAAAATCTAGTGGCCAGCGCGCCAGTACCGGTAGCACTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATCTAGTGGCCAGCGCGCCAGTACCGGTAGCACTTCGGGACTACGCT 3300

RESULT 2

US-09-765-555-15
; Sequence 15, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014, 40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger
; OTHER INFORMATION: protein ZFPm2
US-09-765-555-15

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGGCCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGGCCCGGAGAGA 60
Qy 61 GTCATTCAGGGTGTGAATGTGAACCAAGTAACGTTATACGATGTCGAGAGTATGCG 120
Db 61 GTCATTCAGGGTGTGAATGTGAACCAAGTAACGTTATACGATGTCGAGAGTATGCG 120
Qy 121 GTGTCCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCACCGCTTTCTGCGAAA 180
Db 121 GTGTCCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCACCGCTTTCTGCGAAA 180
Qy 181 CCGGGAAGAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCACCAACCGCTGGCAC 240
Db 181 CCGGGAAGAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCACCAACCGCTGGCAC 240
Qy 241 AACAACTGGCGGGCAACAGTGTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAACAGTGTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGGCGCTCGCAATTCGCGCGGATTAATCTCGCGCGCATCAATCTGGGTGGCCAGC 360
Db 301 ACGGCGCTCGCAATTCGCGCGGATTAATCTCGCGCGCATCAATCTGGGTGGCCAGC 360
Qy 361 TGGTGTGCTGATGTAGAACGAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAAAATC 420
Db 361 TGGTGTGCTGATGTAGAACGAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAAAATC 420
Qy 421 TTCTCGCCAAACGCGTCAAGTGGGTGATCAATTAATCTATCCGTGATGACAGGATGCCA 480
Db 421 TTCTCGCCAAACGCGTCAAGTGGGTGATCAATTAATCTATCCGTGATGACAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTATGTTCCGGCGTATTTCTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTATGTTCCGGCGTATTTCTGATGCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTACCGCACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTACCGCACTGGCGGTGGAGCATC 600
Qy 601 TGGTCCCATTTGGGTCAACCAAGCAATTCGCGCTGTTAGCGGGGCCCATTTAAAGTTCTGTCGG 660

Db	601	TGGT	CGCAAT	GGGT	CACAG	CAAAAT	CGCG	CTGTT	TAG	CGGG	CCCAAT	TAAG	TTCT	GTCT	CGG	660							
Qy	661	CGCG	TT	CGGT	CT	GGCT	TGG	CAATAA	ATAT	CT	CAC	TCG	CAAT	CAAA	ATTCAG	CGCGATAG	720						
Db	661	CGCG	TT	CGGT	CT	GGCT	TGG	CAATAA	ATAT	CT	CAC	TCG	CAAT	CAAA	ATTCAG	CGCGATAG	720						
Qy	721	CGGA	ACGG	GAAG	CGCA	TG	CGAG	TGCC	AT	TG	CCG	TTT	TC	AC	AAAC	CAATG	CAAAATGCTGA	780					
Db	721	CGGA	ACGG	GAAG	CGCA	TG	CGAG	TGCC	AT	TG	CCG	TTT	TC	AC	AAAC	CAATG	CAAAATGCTGA	780					
Qy	781	ATG	AGGG	CAT	CGT	TCC	CACT	CGCAT	TG	CTG	TTG	CC	AC	AGAT	CG	AT	CGCGCGCA	840					
Db	781	ATG	AGGG	CAT	CGT	TCC	CACT	CGCAT	TG	CTG	TTG	CC	AC	AGAT	CG	AT	CGCGCGCA	840					
Qy	841	TG	CGCG	CCAT	TAC	CGAG	TC	CGG	CT	CG	CGT	TG	TG	CGGA	TAT	CT	CGGT	AGTGGGATACG	900				
Db	841	TG	CGCG	CCAT	TAC	CGAG	TC	CGG	CT	CG	CGT	TG	TG	CGGA	TAT	CT	CGGT	AGTGGGATACG	900				
Qy	901	AC	GAT	AC	CGA	AG	CAG	CT	CAT	TG	TAT	AT	CC	CG	CGT	TAA	CAC	CAGAT	TTTC	960			
Db	901	AC	GAT	AC	CGA	AG	CAG	CT	CAT	TG	TAT	AT	CC	CG	CGT	TAA	CAC	CAGAT	TTTC	960			
Qy	961	GC	CT	GT	TG	GGG	CAAA	CC	CAG	CGT	GA	CCG	TT	TG	CT	CA	CT	CAG	CGCCAG	CGGTGA	1020		
Db	961	GC	CT	GT	TG	GGG	CAAA	CC	CAG	CGT	GA	CCG	TT	TG	CT	CA	CT	CAG	CGCCAG	CGGTGA	1020		
Qy	1021	AG	GC	CAAT	CAG	CT	CT	CG	CGT	CT	C	ACT	GT	GT	GA	AA	AA	AA	AA	AA	1080		
Db	1021	AG	GC	CAAT	CAG	CT	CT	CG	CGT	CT	C	ACT	GT	GT	GA	AA	AA	AA	AA	AA	1080		
Qy	1081	CG	CAAA	CCG	CT	CT	CC	CG	CGT	TT	G	CC	CA	T	CA	T	TA	T	GC	AG	CT	1140	
Db	1081	CG	CAAA	CCG	CT	CT	CC	CG	CGT	TT	G	CC	CA	T	CA	T	TA	T	GC	AG	CT	1140	
Qy	1141	CC	CG	ACT	G	CA	AA	AG	CG	CG	CAG	TG	CG	CA	A	CG	CA	A	TA	T	TA	1200	
Db	1141	CC	CG	ACT	G	CA	AA	AG	CG	CG	CAG	TG	CG	CA	A	CG	CA	A	TA	T	TA	1200	
Qy	1201	GC	CAAT	CT	CAT	CT	TT	TG	CA	AG	CT	TAT	AT	C	AG	CT	GC	AG	GT	GC	CA	1260	
Db	1201	GC	CAAT	CT	CAT	CT	TT	TG	CA	AG	CT	TAT	AT	C	AG	CT	GC	AG	GT	GC	CA	1260	
Qy	1261	TC	AG	CG	AG	CCAT	CG	GA	AG	CT	TG	GT	GA	AG	CT	GC	AG	GT	GC	GA	TA	1320	
Db	1261	TC	AG	CG	AG	CCAT	CG	GA	AG	CT	TG	GT	GA	AG	CT	GC	AG	GT	GC	GA	TA	1320	
Qy	1321	TG	TC	GC	T	CA	AG	GG	CA	CT	CC	CG	TT	CT	CG	TA	A	T	T	T	T	1380	
Db	1321	TG	TC	GC	T	CA	AG	GG	CA	CT	CC	CG	TT	CT	CG	TA	A	T	T	T	T	1380	
Qy	1381	CT	GG	CA	AA	TAT	CT	G	AA	AT	G	AG	CT	TT	TG	CA	A	T	T	T	T	1440	
Db	1381	CT	GG	CA	AA	TAT	CT	G	AA	AT	G	AG	CT	TT	TG	CA	A	T	T	T	T	1440	
Qy	1441	ATT	GT	G	AG	CG	GA	TAA	CA	AA	TT	TC	AC	AG	GA	AA	CAG	CG	AG	T	CG	1500	
Db	1441	ATT	GT	G	AG	CG	GA	TAA	CA	AA	TT	TC	AC	AG	GA	AA	CAG	CG	AG	T	CG	1500	
Qy	1501	G	C	ACT	T	C	AC	CA	A	A	G	G	C	A	T	A	T	T	G	A	A	A	1560
Db	1501	G	C	ACT	T	C	AC	CA	A	A	G	G	C	A	T	A	T	T	G	A	A	A	1560
Qy	1561	A	T	T	A	A	A	A	A	A	G	G	C	T	A	A	T	T	A	A	A	A	1620
Db	1561	A	T	T	A	A	A	A	A	A	G	G	C	T	A	A	T	T	A	A	A	A	1620
Qy	1621	A	C	C	G	G	A	A	T	T	A	A	A	G	T	T	A	A	A	A	A	A	1680
Db	1621	A	C	C	G	G	A	A	T	T	A	A	A	G	T	T	A	A	A	A	A	A	1680
Qy	1681	G	C	G	G	C	A	A	C	T	G	G	C	C	T	G	A	C	A	T	T	C	1740

Db	1681	CGCGCAACTGCGGATGGCCCTGACAATTATCTTCTGGGCACACGACCGCTTGTGGTGCGTAC	1740
Qy	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGACACAAGCGTTCCAGGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGACACAAGCGTTCCAGGACAAGCTGTAT	1800
Qy	1801	CGGTTTACCTGGGATGCGGTACGTTTACAACGCGCAAGCTGATTTGCTTACC CGATCGCTGTT	1860
Db	1801	CGGTTTACCTGGGATGCGGTACGTTTACAACGCGCAAGCTGATTTGCTTACC CGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTTATACAAGATCTGCTGCGCAACCGCCCAAAAAA CTTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATACAAGATCTGCTGCGCAACCGCCCAAAAAA CTTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATTAAGAACTCAAAAGCGAAAGTAAAGCGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATTAAGAACTCAAAAGCGAAAGTAAAGCGCGCTGATGTTTCAAC	1980
Qy	1981	CTGCAAGNAACGTTACTTCACTGCGCCGCTGATTTGCTCTGACGGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGNAACGTTACTTCACTGCGCCGCTGATTTGCTCTGACGGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAAAAGCGCAAGTACGACATTAAAGACGTTGGGCGTGATTAACGCTGCGCGGAAGCG	2100
Db	2041	TATGAAAAAGCGCAAGTACGACATTAAAGACGTTGGGCGTGATTAACGCTGCGCGGAAGCG	2100
Qy	2101	GGTCTGACTTCTGTTGTGACCTGATTTAAAAAACAACACATGAATGAGACACCGGATTAC	2160
Db	2101	GGTCTGACTTCTGTTGTGACCTGATTTAAAAAACAACACATGAATGAGACACCGGATTAC	2160
Qy	2161	TCCATCGCAGAAGCTGCTTTTAAATAAGGCGGAACAGCGATGCCATCAACGCGCCGTGG	2220
Db	2161	TCCATCGCAGAAGCTGCTTTTAAATAAGGCGGAACAGCGATGCCATCAACGCGCCGTGG	2220
Qy	2221	GCATPGTCCAACTCGACACCGCAAGTGAATTTATGGTTAAACGGTACTTGC CGACCTTC	2280
Db	2221	GCATPGTCCAACTCGACACCGCAAGTGAATTTATGGTTAAACGGTACTTGC CGACCTTC	2280
Qy	2281	AAGGCTCAACCATCCAAACCGTTTGGTGGCGTCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Db	2281	AAGGCTCAACCATCCAAACCGTTTGGTGGCGTCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Qy	2341	CCGAAACAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAAACAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
Qy	2401	GAAGCGTTTAAATAAGNAACCGCTGGTGGCGTACGCGTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGTTTAAATAAGNAACCGCTGGTGGCGTACGCGTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCAAAAGATCCAGTATTGCGCGCACCATGGAAAAACGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCAAAAGATCCAGTATTGCGCGCACCATGGAAAAACGCCAGAAAGGTGAATCATG	2520
Qy	2521	CCGAAACATCCCGCAGATGTCGCTTTTGTGTATGCGGTGCGTATCTGCGGTATCAACGCC	2580
Db	2521	CCGAAACATCCCGCAGATGTCGCTTTTGTGTATGCGGTGCGTATCTGCGGTATCAACGCC	2580
Qy	2581	GCAGCGGTCTGACTGTGATGAAGCCCTGAAGACGCCGACACTAATTCGAGCTCG	2640
Db	2581	GCAGCGGTCTGACTGTGATGAAGCCCTGAAGACGCCGACACTAATTCGAGCTCG	2640
Qy	2641	AACAACAACAATAAACAATAACAACAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC	2700
Db	2641	AACAACAACAATAAACAATAACAACAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC	2700
Qy	2701	GGATCTCTTCTCTGTGCCCGAGCGGCTCGAGCCGGGGAGAGCCCTATGCTTGT	2760
Db	2701	GGATCTCTTCTCTGTGCCCGAGCGGCTCGAGCCGGGGAGAGCCCTATGCTTGT	2760
Qy	2761	CCGGAATGTGGTAACTCTTCTCTCAGAGCTCTCACTGCTGCGCCACCAAGCGGTACCCAC	2820
Db	2761	CCGGAATGTGGTAACTCTTCTCTCAGAGCTCTCACTGCTGCGCCACCAAGCGGTACCCAC	2820

```
QY 2821 ACGGTGAAAAACCGTATATAATGCCAGTCCGGCAAAATCTTTTAGCGAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATATAATGCCAGTCCGGCAAAATCTTTTAGCGAGTCCAGCAAC 2880
QY 2881 CTGGTGCCTCAACAGCACTCATCTGCGAGAGAGCCATACAAATGTCCAGAAATGTGC 2940
Db 2881 CTGGTGCCTCAACAGCACTCATCTGCGAGAGAGCCATACAAATGTCCAGAAATGTGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTCACACCGGGGAGAAG 3000
QY 3001 CCCTATGTTGTCGGGAATGTGGTAAGTCTTTTACGCGCAGGAGTAACCTGTGTGCGCCAC 3060
Db 3001 CCCTATGTTGTCGGGAATGTGGTAAGTCTTTTACGCGCAGGAGTAACCTGTGTGCGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATATAATGCCAGAGTCCGGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATATAATGCCAGAGTCCGGCAAAATCTTTTAGC 3120
QY 3121 CAGCGCGGCCACTCGCCAGCCATCAAGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGGCCACTCGCCAGCCATCAAGCACTCATCTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTAC 3240
QY 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGACAGTACCGGTACAGGTTCCGGACCTACCGT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGACAGTACCGGTACAGGTTCCGGACCTACCGT 3300
```

RESULT 3

```
US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-Ap3 and zinc finger
; OTHER INFORMATION: protein ZFPap3
US-09-765-555-18
```

```
Query Match 97.8%; Score 3226.4; DB 10; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 CCGACACCATCAATGGTGCAGAAACCTTTTCGGGTATGCGATGATAGCCCGGAGAGA 60
Db 1 CCGACACCATCAATGGTGCAGAAACCTTTTCGGGTATGCGATGATAGCCCGGAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCACGCTTCTGCGAAAA 180
```

```
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGCTTCTGCGAAAA 180
QY 181 CCGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGCGATGCGGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGCGCGGCAAAACAGTCTGTTGATTTGGGTTGCCACCTTCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGCGGCAAAACAGTCTGTTGATTTGGGTTGCCACCTTCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCCAAAATGTCGCGCGGATTAATCTCGCGCGCATCAATCTGGGTGCGCAGCG 360
Db 301 ACGCGCGTCCCAAAATGTCGCGCGGATTAATCTCGCGCGCATCAATCTGGGTGCGCAGCG 360
QY 361 TGGTGGTTCGATGATGAACGAGCGCGTTCGAAGCCTCTAAAGCGCGGTGACAAATC 420
Db 361 TGGTGGTTCGATGATGAACGAGCGCGTTCGAAGCCTCTAAAGCGCGGTGACAAATC 420
QY 421 TTCTCGCGCAACGCGTCACTGATGATTAATCTAACTATCCGCTGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCACTGATGATTAATCTAACTATCCGCTGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTAATCTTCTGATCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTAATCTTCTGATCTCTGACCCAGA 540
QY 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTCTGTCTGG 660
Db 601 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTCTGTCTGG 660
QY 661 CGCGTCTCGGTCTGGCTGGCGATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGGTCTGGCTGGCGATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAAACGGGAAAGGACCTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAAACGGGAAAGGACCTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCTGTTCCCACTGCGATGTGGTTCGCAACGATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGCATCTGTTCCCACTGCGATGTGGTTCGCAACGATCAGATGCGCTGGCGGCA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGTGGCGGTGGTGGGATATCTCGGTAGTGGGATAGC 900
Db 841 TCGCGGCCATTACCGAGTCCGGGTGGCGGTGGTGGGATATCTCGGTAGTGGGATAGC 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCGCGTCTCACTGTTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGTTGCGCGTCTCACTGTTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCAATTAATGAGCTGGCGACGACAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCAATTAATGAGCTGGCGACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTACTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTACTTAG 1200
QY 1201 GCACAATTTCTATGTTTACAGCTTATCATCGACTGCGGTGACCAATGCTTCTGGCG 1260
```

Db 1201 GCACAAATTCATGTTGACAGCTTATCATCOAGTGCACGGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGCAGCATCGGAAGCTGTGGTATAGCTGTGAGGTCGTAATCACTGCATCAATTCG 1320
Db 1261 TCAGGCAGCATCGGAAGCTGTGGTATAGCTGTGAGGTCGTAATCACTGCATCAATTCG 1320
Qy 1321 TGTGCGTCAAGCGGCACTCCGCTTCTGGATAATGTTTTTGGCCGACATCAATACGGTT 1380
Db 1321 TGTGCGTCAAGCGGCACTCCGCTTCTGGATAATGTTTTTGGCCGACATCAATACGGTT 1380
Qy 1381 CTGGCAAAATATCTGAAATGAGCTGTGACATTAATCATCGGCTGTAATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTGACATTAATCATCGGCTGTAATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTAGTGTTTTTACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTAGTGTTTTTACGA 1500
Qy 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACCTGAAGAGGTAATCTGTAATCTGG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACCTGAAGAGGTAATCTGTAATCTGG 1560
Qy 1561 ATTAACGCGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGCGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 GCGCAATCGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGTAC 1740
Db 1681 GCGCAATCGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATACACCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATACACCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGTACGTTTACACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGTACGTTTACACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGTCGGAAACCCGCAAAACCTCGGAA 1920
Db 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGTCGGAAACCCGCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGGCGTGGATAAAGAACTGAAAGCGAAAGTAAGACGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGTGGATAAAGAACTGAAAGCGAAAGTAAGACGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTGCTGTACGCGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTGCTGTACGCGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGATTAAGACGTCGGCGTGGATACGTCGGCGGCAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGATTAAGACGTCGGCGTGGATACGTCGGCGGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGNAGTGGCTTTTAAAGGGGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCGCAGNAGTGGCTTTTAAAGGGGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTTATGTTGTAAGCTGACCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTTATGTTGTAAGCTGACCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340

RESULT 4

US-09-765-555-17
; Sequence 17, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate

Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCCCTGGTGGCTAGCGTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCCCTGGTGGCTAGCGTGAAGTCTTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAAGATCCAGTATTTGCGCCACCATGGAACCCCAAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAAGATCCAGTATTTGCGCCACCATGGAACCCCAAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGATGCGTGGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGATGCGTGGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGCTCAGACTGTGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTCAGACTGTGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
Qy 2641 AACCAACAAACAAATAACAATAACAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATAACAATAACAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCTTCTCTGTGGCCAGGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACTGCTGGCGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTCTCAGCAGAGCAGCTCCTCTGCTGGCCACACAGCGTACCCAC 2820
Qy 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTACGCAAGTCCAGCAAC 2880
Db 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTACGCAAGTCCAGCAAC 2880
Qy 2881 CTGCTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTTGGC 2940
Db 2881 CTGCTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTTGGC 2940
Qy 2941 AAGTCTTCTCTCGGTCTGACAAATCTCTCGCCGACCAACAGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTCTCGAGCAGTCCAGCAACCTGGTGGCGCCACCAACGTAATCACACCGGGGAGAAG 3000
Qy 3001 CCTATGCTGTGTCGGAATGTTGTAAGTCTTCTCAGCGCGAGCGATAACTGTTGGCGCCAC 3060
Db 3001 CCTATGCTGTGTCGGAATGTTGTAAGTCTTCTCAGCAACCGAGTGGCTCTCTTGGTTAGACAC 3060
Qy 3061 CAGCGTACCCACACCGGGTGAATAAATGCCAGAGTGGCGCAAAATCTTTTATG 3120
Db 3061 CAGCGTACCCACACCGGGTGAATAAATGCCAGAGTGGCGCAAAATCTTTTATG 3120
Qy 3121 CAGCGCGCCACCTGGAACGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTGGAACGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Qy 3181 CCAGAATGTTGGCAAGTCTTCTCTCGTCTGCAATCTCTGCGCGGACCAACGTAATCTAC 3240
Db 3181 CCAGAATGTTGGCAAGTCTTCTCTCACTTCAAGGCACTTGGTCCGTCAACCAAGTACTAC 3240
Qy 3241 ACCGGTAAAAAACTAGTGGCCAGGCGCGCCAGTACCCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGGCGCGCCAGTACCCGTACGACGTTCCGGACTACGCT 3300

; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein 2PPm4
US-09-765-555-17

Query Match 97.4%; Score 3215.2; DB 10; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAACTTTTCGGGTATGGCATAGAGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAACTTTTCGGGTATGGCATAGAGCCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAACCGTTATACGATGTCGAGAGTATGCG 120
DB 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAACCGTTATACGATGTCGAGAGTATGCG 120
QY 121 GTGTCTTATCAGACCGTTTCCTCCGGTGGTGAACACGAGCCAGCCAGTTCCTCGAAAA 180
DB 121 GTGTCTTATCAGACCGTTTCCTCCGGTGGTGAACACGAGCCAGCCAGTTCCTCGAAAA 180
QY 181 CCGGGGAAAGTGAAGCGGCGATGGCGAGCTGAATTACATTCCTCCACCGCTGGCAC 240
DB 181 CCGGGGAAAGTGAAGCGGCGATGGCGAGCTGAATTACATTCCTCCACCGCTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTGGTTCGATGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTGGTTCGATGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 AGCGCCGTCGCAAAATGTTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCAGCG 360
DB 301 AGCGCCGTCGCAAAATGTTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCAGCG 360
QY 361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC 420
DB 361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGGATGCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGGATGCA 480
QY 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCGGCGGTTATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCGGCGGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGGCTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGGCTGGCGGTGGAGCATC 600
QY 601 TGGTTCGATTTGGTTCACAGCAAACTCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTTCGATTTGGTTCACAGCAAACTCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CCGGCTCGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CCGGCTCGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAAGCGGAGCGGATGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CCGAAGCGGAGCGGATGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780

DB 721 CCGAAGCGGAGCGGATGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCCCATCGCATGCTGGTGGCAACGATCAGATGCGCTGGCGCGAA 840
DB 781 ATGAGGGCATCGTTCCTCCCATCGCATGCTGGTGGCAACGATCAGATGCGCTGGCGCGAA 840
QY 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGCAAAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCTCTCAGTGGTGAAGAAAGAAACCACTTGGCGGCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCTCTCAGTGGTGAAGAAAGAAACCACTTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGTGGCAGACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGTGGCAGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
QY 1201 GCACAAATTCATGTTTGACAGCTTATCATCGATGACCGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATTCATGTTTGACAGCTTATCATCGATGACCGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1320
DB 1261 TCAGCAGCCATCGGAAGCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1320
QY 1321 TGTGCTCAAGCGGCACATCCCGCTTCTGGATTAATGTTTTTGGCGCGACATCATACGGTT 1380
DB 1321 TGTGCTCAAGCGGCACATCCCGCTTCTGGATTAATGTTTTTGGCGCGACATCATACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTAAATGATGTGGA 1440
DB 1381 CTGGCAAAATATCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTAAATGATGTGGA 1440
QY 1441 ATTGTGAGCGGATTAACAAATTTACACAGCAACGAGTCCGTTTAGTGTGTTTTCACGA 1500
DB 1441 ATTGTGAGCGGATTAACAAATTTACACAGCAACGAGTCCGTTTAGTGTGTTTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTACCGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGCAACTGCGGATGGGCTGACATTAATCTTGGGCAACGACCGCTTGGTGGCTAC 1740
DB 1681 GCGCAACTGCGGATGGGCTGACATTAATCTTGGGCAACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATTCGCGCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGAAGAAGTGTAT 1800
DB 1741 GCTCAATTCGCGCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGAAGAAGTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACCGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
DB 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACCGCAAGCTGATTGCTTACCCGATCCCTGTT 1860

1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGTCGGAACCCGCCAAAACCTGGGAA 1920
1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGTCGGAACCCGCCAAAACCTGGGAA 1920
1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAGGTAGAGCGCGCTGATGTTCAAC 1980
1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAGGTAGAGCGCGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATGTCGTCGACGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATGTCGTCGACGGGGTTATGCGTTCAAG 2040
2041 TATGAAAAACGGAAGTACGACATTAAGACGCTGGCGTGGATAAGCTGGCGGCGAAGCG 2100
2041 TATGAAAAACGGAAGTACGACATTAAGACGCTGGCGTGGATAAGCTGGCGGCGAAGCG 2100
2101 GGTCTGACCTTCCTGGTTGACCTGATTAAGGCGAAGATGACCATCAACGGCCCGTGG 2160
2101 GGTCTGACCTTCCTGGTTGACCTGATTAAGGCGAAGATGACCATCAACGGCCCGTGG 2160
2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAAGATGACCATCAACGGCCCGTGG 2220
2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAAGATGACCATCAACGGCCCGTGG 2220
2221 GCATGGTCCAAATCGACACGAGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
2221 GCATGGTCCAAATCGACACGAGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
2281 AAGGGTCAACCATCAACACCGTTCGTTGGCGTGTGAGCGAGGTATTAAGCGCGCAGT 2340
2281 AAGGGTCAACCATCAACACCGTTCGTTGGCGTGTGAGCGAGGTATTAAGCGCGCAGT 2340
2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
2401 GAAGCGGTTAATAAGCAAAACCGCTGGTTCGCTGAGCTGAACTCTTACGAGGAAGAG 2460
2401 GAAGCGGTTAATAAGCAAAACCGCTGGTTCGCTGAGCTGAACTCTTACGAGGAAGAG 2460
2461 TTGGCGAAGATCCACGATTTGCGCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
2461 TTGGCGAAGATCCACGATTTGCGCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
2521 CCGAACATCCCGCAGATGTCGCTTTCTGATGCGGTGCTGCTGCTGATCAACGCC 2580
2521 CCGAACATCCCGCAGATGTCGCTTTCTGATGCGGTGCTGCTGCTGATCAACGCC 2580
2581 GCCAGCGTCTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
2581 GCCAGCGTCTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
2641 AACAAACAACAATAACAATAACAACCTCGGATCGGAGGAAGATTCAGAAATC 2700
2641 AACAAACAACAATAACAATAACAACCTCGGATCGGAGGAAGATTCAGAAATC 2700
2701 GGATCCTCTTCTCTGTCGCGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
2701 GGATCCTCTTCTCTGTCGCGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACTGGTGGCCCAACGAGCTACCCAC 2820
2761 CCGGAATGTGTAAGTCTTCTCAGAGCTCTCACTGGTGGCCCAACGAGCTACCCAC 2820
2821 ACGGGTGAAGAACCGTAAATGCCAGAGTGGGCAAACTCTTTAGCCAGTCCAGCAAC 2880
2821 ACGGGTGAAGAACCGTAAATGCCAGAGTGGGCAAACTCTTTAGCCAGTCCAGCAAC 2880
2881 CTGTCGCGCCATCAACGACTCATATCTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
2881 CTGTCGCGCCATCAACGACTCATATCTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940

2941 AAGTCTTTCTCTCGTCTGACAAATCTCGTCCGGCACCAAGTACTCACCAGGGGAGAAG 3000
2941 AAGTCTTTCTCTCGTCTGACAAATCTCGTCCGGCACCAAGTACTCACCAGGGGAGAAG 3000
3001 CCTATGCTTTGTCGGAATGTGGTAACTCTTTCAGCGCAGAGATAACCTGGTGGCCAC 3060
3001 CCTATGCTTTGTCGGAATGTGGTAACTCTTCTCTCAGAGCTCTCACCTGGTGGCCAC 3060
3061 CAGCGTACCACACGCGGTGAAGAACCGTATAAATGCCAGAGTGGCGCAATCTTTTACG 3120
3061 CAGCGTACCACACGCGGTGAAGAACCGTATAAATGCCAGAGTGGCGCAATCTTTTACG 3120
3121 CAGCGCGCCACCTGGCGCAGCATCAACGCACTCATATCTGGCGAGAGCCATACAAATGT 3180
3121 CAGCGCGCATCAACCTGGTGGCGCATCAACGCACTCATATCTGGCGAGAGCCATACAAATGT 3180
3181 CAGAGATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAAGTACTCAC 3240
3181 CAGAGATGTGGCAAGTCTTTCTCAACTTCAGGCGCATTTGGTCCGTCAACCAACGTACTCAC 3240
3241 ACCGTTAAAAAACTAGTGGCGCGCCAGTACCCGTACGACGTTCCGGACTACGCT 3300
3241 ACCGTTAAAAAACTAGTGGCGCGCCAGTACCCGTACGACGTTCCGGACTACGCT 3300

RESULT 5
US-09-765-555-16
; Sequence 16, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; CURRENT APPLICATION NUMBER: US/09/765,555
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m3 and zinc finger
; OTHER INFORMATION: protein ZFPm3
US-09-765-555-16

Query Match 97.2%; Score 3207.2; DB 10; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGTACGCGCCCGAAGAGA 60
1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGTACGCGCCCGAAGAGA 60
61 GTCAATTCAGGGTGGTGAATGTGAAAACCGATTAACGATGTGCGAGATGTGCGG 120
61 GTCAATTCAGGGTGGTGAATGTGAAAACCGATTAACGATGTGCGAGATGTGCGG 120
121 GTGCTCTTATCAGCCGTTTCCGGTGTGAACCCAGGCCAGCCAGCTTCTTCCGAAA 180
121 GTGCTCTTATCAGCCGTTTCCGGTGTGAACCCAGGCCAGCCAGCTTCTTCCGAAA 180
181 CCGGGGAAAAGTGGAGCGCGCATGGCGAGCTGAATTAATTTCCCAACCGCGTGGCAC 240
181 CCGGGGAAAAGTGGAGCGCGCATGGCGAGCTGAATTAATTTCCCAACCGCGTGGCAC 240
241 AACCACTGGCGGCAACAGTCTGCTGATTTGGCGTGGCCACTTCAGTCTGGCCCTGC 300

Db 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCGCGCCCTGC 300
Qy 301 ACGGCCCTCGCAAAATGTCGCGCGAATTAATCTCGCGCGGATCAACTGGGTGCCAGC 360
Db 301 ACGGCCCTCGCAAAATGTCGCGCGAATTAATCTCGCGCGGATCAACTGGGTGCCAGC 360
Qy 361 TGGTGGTCTGATGATGAGAAAGAGCGCGTCGAGCGCTGAAAGCGCGGTGACCAATC 420
Db 361 TGGTGGTCTGATGATGAGAAAGAGCGCGTCGAGCGCTGAAAGCGCGGTGACCAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACATATCCGCTGGATCACCAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACATATCCGCTGGATCACCAGGATGCCA 480
Qy 481 TTGCTGTGGAAAGCTGCCCTGACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACAGA 540
Db 481 TTGCTGTGGAAAGCTGCCCTGACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACAGA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAGACGGTACGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAGACGGTACGGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGGCATGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCGCG 660
Db 601 TGGTGGCATGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCGCG 660
Qy 661 CGCGTCTCGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGGGATGCTGGTGGCCAAACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGGGATGCTGGTGGCCAAACGATCAGATGGCGTGGCGCAA 840
Qy 841 TGCGCGCATTAACGAGTCCGGGTCGCGGTGGTGGCGATATCTCGTAGTGGGATACG 900
Db 841 TGCGCGCATTAACGAGTCCGGGTCGCGGTGGTGGCGATATCTCGTAGTGGGATACG 900
Qy 901 ACGATACGGAAGACATGTTATATCCGCGGTTAACCCCATCAAAACGATTTTC 960
Db 901 ACGATACGGAAGACATGTTATATCCGCGGTTAACCCCATCAAAACGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGCAACGAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGCAACGAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
Qy 1201 GCACAATCTCATGTTTGCACGCTTATCATCGACTGCAGGTGCAACATGCTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGCACGCTTATCATCGACTGCAGGTGCAACATGCTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGGTATGAGCTGTGAGGTCGTAATCACTGCAATATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGAGCTGTGAGGTCGTAATCACTGCAATATTCG 1320
Qy 1321 TGTCTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTCTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380

Qy 1381 CTGCAAAATATCTTGAAATGAGCTGTTGAACAATTAATCATTCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGCAAAATATCTTGAAATGAGCTGTTGAACAATTAATCATTCGGCTCGTATTAATGTGTGA 1440
Qy 1441 ATTCTGAGCGGATAACAATTTTACACAGGAACACGAGTCGCTTTAGTGTGTTTTCAGCA 1500
Db 1441 ATTCTGAGCGGATAACAATTTTACACAGGAACACGAGTCGCTTTAGTGTGTTTTCAGCA 1500
Qy 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTCGTAACTCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTCGTAACTCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGCTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGCTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCGCCACAGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGGAACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGGAACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGTACGTTTCAACCGCAAGCTGATGTTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTTCAACCGCAAGCTGATGTTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAACGGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAACGGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATCAAAAACGCAAGTACGACATTAAGACCGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATCAAAAACGCAAGTACGACATTAAGACCGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACAACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAATAAGGCGCAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGGCGCAACAGCGATGACCATCAAGCGCCGCTGG 2220
Qy 2221 GCATGGTCAACATCGACACCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGAGGTTAATGAGCCGCCAGT 2340
Db 2281 AAGGTTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGAGGTTAATGAGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCCGAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACCTATCTGCTGACTGATGAAGGTTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAGCTTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAGCTTTTACGAGGAAGAG 2460


```
QY 2461 TTGGCGAAAGATCCAGTATTGCGCCACCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGCCACCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAATATCCCGAGATGTCGGCTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAATATCCCGAGATGTCGGCTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGTCAGACTGTTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGACTGTTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACAAACAATAAACAATAACAACAACTCGGGATCGAGGGAAGGATTCGAAATTC 2700
Db 2641 AACACAAACAATAAACAATAACAACAACTCGGGATCGAGGGAAGGATTCGAAATTC 2700
QY 2701 GGATCCTCTTCTCTGTCGGCCAGCGGCCCTCGAGCCCGGGGAGAAGCCCTATGCTTGT 2760
Db 2701 GGATCCTCTTCTCTGTCGGCCAGCGGCCCTCGAGCCCGGGGAGAAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCTGTGCGCCACCCAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTCTCAGAGCTCTCACCTGTGCGCCACCCAGCGTACCCAC 2820
QY 2821 ACGGTGAAAGACGTATATAATGCCAGAGTCGCGCAAACTCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAGACGTATATAATGCCAGAGTCGCGCAAACTCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGCCTCAACACGCACTCATACTGCGAGAGCCATACAAATGTCAGAGATGGC 2940
Db 2881 CTGGTGCCTCAACACGCACTCATACTGCGAGAGCCATACAAATGTCAGAGATGGC 2940
QY 2941 AAGTCTTCTCTGCTGTCGATCAATCTGTCGCGCACCAACAGTACTCACCGGGAGAAG 3000
Db 2941 AAGTCTTCTCAGCCAGAGCTCCAGCTGTGTCGCGCACCAACAGTACTCACCGGGAGAAG 3000
QY 3001 CCCTATGCTGTCGCGAATGTTAAGTCTTCTGAGCCGCGAGTACCGCGAGTACCTGTCGCCAC 3060
Db 3001 CCCTATGCTGTCGCGAATGTTAAGTCTTCTGAGCCGCGAGTACCTGTCGCCAC 3060
QY 3061 CAGCGTACCCACACGGGTGAAAGACCGTATATAATGCGCAGAGTCGCGCAAACTCTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAGACCGTATATAATGCGCAGAGTCGCGCAAACTCTTTAGT 3120
QY 3121 CAGCGCGCACCTGCGCAGCCATCAACGCACTCATATGCGGAGAGCCATACAAATGT 3180
Db 3121 GACTGCGCGCACTTCTCTCGGTCTGACAACTCAACGCACTCATATGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGACAACTCTGTCGCGCACCAACGTAATCTAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGACAACTCTGTCGCGCACCAACGTAATCTAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGAGTACCGGTACGACGTTCCGGGACTAGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGAGTACCGGTACGACGTTCCGGGACTAGCT 3300
```

RESULT 6

```
US-10-257-384A-1
; Sequence 1, Application US/10257384A
; Publication No. US20040087524A1
; GENERAL INFORMATION:
; APPLICANT: Wiedersanders, Bernd
; APPLICANT: Maubach, Gunter
; TITLE OF INVENTION: Agent for postoperative use after removal of bone tumors
; FILE REFERENCE: 2945-101
; CURRENT APPLICATION NUMBER: US/10/257,384A
; CURRENT FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: PCT/DE 01/01510
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: DE 100 20 125.3
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 19
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7373
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Expression vector pMalc2cchbmp2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)..(3453)
; OTHER INFORMATION: /product = fusion protein MBP Cystatin C BMP-2
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2689)
; OTHER INFORMATION: Faktor Xa Protease /position=388
; OTHER INFORMATION: (Amino acid sequence)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3067)..(3108)
; OTHER INFORMATION: /spacer between Cystatin C andBMP-2
; PUBLICATION INFORMATION:
; AUTHORS: Zwick, M B.
; TITLE: Expression vector pMal-X, complete sequence
; JOURNAL: Anal. Biochem.
; VOLUME: 264
; ISSUE: 1
; PAGES: 87-97
; DATE: 1998-11-01
; DATABASE ACCESSION NUMBER: Entrez Nucleotide database/AF031813
; DATABASE ENTRY DATE: 2001-05-07
; RELEVANT RESIDUES: 1528 - 3453
US-10-257-384A-1
```

```
Query Match 82.0%; Score 2705.8; DB 18; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 CCAGACCACTCGAATGGTGCAGAAACCTTTCCGGGTATGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCAGACCACTCGAATGGTGCAGAAACCTTTCCGGGTATGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGGTGAATGTGAACACCAAGTAACTATACATGTCGAGAGTATGCCG 120
Db 61 GTCATTCAGGGTGGTGAATGTGAACACCAAGTAACTATACATGTCGAGAGTATGCCG 120
QY 121 GTGTCTTTATCAGACCGTTTCCCGGTGTGAACACCAAGTAACTATACATGTCGAGAGTATGCCG 180
Db 121 GTGTCTTTATCAGACCGTTTCCCGGTGTGAACACCAAGTAACTATACATGTCGAGAGTATGCCG 180
QY 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATTTCTCGGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATTTCTCGGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGTCGATGTGAGAACGAAAGCGGCTGTAAGCGCGGTGTAAGCGGGTGCACATC 420
Db 361 TGGTGTGTGTCGATGTGAGAACGAAAGCGGCTGTAAGCGCGGTGTAAGCGGGTGCACATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATTCCTCCGCTGGATGACAGAGTGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATTCCTCCGCTGGATGACAGAGTGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCACATAATTTTCCGGCGTTATTTCTTGATGTCTTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACATAATTTTCCGGCGTTATTTCTTGATGTCTTGACCCAGA 540
```


QY 541 CACCCATCAACGATATTATTTCTCCCATGAAGACGGTACCGGACTGGCGCTGGAGCATC 600
Db 541 CACCCATCAACGATATTATTTCTCCCATGAAGACGGTACCGGACTGGCGCTGGAGCATC 600
QY 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTTCAACAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTTCAACAACCAATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGGCA 840
Db 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGGCA 840
QY 841 TGGCGGCCATTACGAGTCCGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTACGAGTCCGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAAGATCAAAACGAGTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAAGATCAAAACGAGTTTC 960
QY 961 GCCTCTGGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAGAAAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAGAAAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCACACAGGTTT 1140
Db 1081 CGCAAAACGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCACACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGACGTGAGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGACGTGAGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGAACGCTTATCATGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGAACGCTTATCATGACTGCACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGACAGTTCGTAATCACTGCATTAATCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGACAGTTCGTAATCACTGCATTAATCG 1320
QY 1321 TGTCTGCTCAAGCGCACTCCCGTCTGGGATAATGTTTTTGGCCGACATCAACGGTT 1380
Db 1321 TGTCTGCTCAAGCGCACTCCCGTCTGGGATAATGTTTTTGGCCGACATCAACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTCAACA 1500
Db 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTCAACA 1500
QY 1501 GCATTTCAACACAGGACATAGATTATGAACCTGAAGAGGTAAACTGCTGATCTCG 1560
Db 1501 GCATTTCAACACAGGACATAGATTATGAACCTGAAGAGGTAAACTGCTGATCTCG 1560
QY 1561 ATTAAACGGCGATAAAGGCTATAACGGTCTCGTGAAGTCCGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAAACGGCGATAAAGGCTATAACGGTCTCGTGAAGTCCGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTTAAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT 1680

Db 1621 ACCGGAATTTAAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGCGCAACTGCGGATGGCCCTTGACATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGCGGATGGCCCTTGACATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGACAAAGCGTTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGACAAAGCGTTTCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAAACCCGCCAAAAAACCCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAAACCCGCCAAAAAACCCTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATAAAGAACTGAAGGAAAGSTAAAGGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAAGAACTGAAGGAAAGSTAAAGGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTTAAAGACGTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTTAAAGACGTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAACAACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAACAACATGAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACCAAGTGAATTTATGGTGAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCAAGTGAATTTATGGTGAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCAACCATCAAAACCGTTCTGGCTGTGAGCGCAGGTATTAAACCGCCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTCTGGCTGTGAGCGCAGGTATTAAACCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGGCAAGATTCACAGTATTGCGCCCAACCATGGAAGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGCAAGATTCACAGTATTGCGCCCAACCATGGAAGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCACGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCACGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATTAACAATAACAACTCCGGATCGAGGGAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAACAATTAACAATAACAACTCCGGATCGAGGGAGGATTTTCAAGATTC 2700
QY 2701 GGATCCTCTTCTGCTGGCCAGCGCGCTTCGAGAGCCGG 2741

Db 2701 GGATCCTCCAGTCCCGCAAGCGCGCGCTGGTGGAGG 2741

RESULT 7

US-10-149-472-5
; Sequence 5, Application US/10149472
; Publication No. US20040029204A1
; GENERAL INFORMATION:
; APPLICANT: COUBIN-GRAMATICA, FRANCOISE
; APPLICANT: DUCOMMUN, BERNARD
; APPLICANT: PREVOST, GREGOIRE
; TITLE OF INVENTION: METHOD FOR OBTAINING HUMAN CDC25 PHOSPHATASES AND METHOD
; FILE REFERENCE: bml-427.065
; CURRENT APPLICATION NUMBER: US/10/149,472
; PRIOR APPLICATION NUMBER: PCT/FR00/03496
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: FR 99/15722
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: FR 00/06883
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: FR 00/12008
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 8101
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: JM109 / pMAL - Hs Cdc25C strain
US-10-149-472-5

Query Match 82.0%; Score 2705.8; DB 17; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGTGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGTGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCGAATTCAGGGTGGTGAATGTGAAACAGTAAACGTTATACGATGTGCGAGAGTATGCCG 120
DB 61 GTCGAATTCAGGGTGGTGAATGTGAAACAGTAAACGTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTCCTCCGCGTGGTGAACCGAGCCAGCCAGCGTTCCTGCGAAA 180
DB 121 GTGTCTCTTATCAGACCGTTCCTCCGCGTGGTGAACCGAGCCAGCCAGCGTTCCTGCGAAA 180
QY 181 CGCGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTCATTCCTCCAAACCGCGTGGCAC 240
DB 181 CGCGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTCATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGAACAGTCGTGTCTGATGGCGTGGCCACTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGAACAGTCGTGTCTGATGGCGTGGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTGGCAAAATGTGCGCGGATTAATCTCTCGCGCGATCAACTGGTGGCAGCG 360
DB 301 ACGCGCGTGGCAAAATGTGCGCGGATTAATCTCTCGCGCGATCAACTGGTGGCAGCG 360
QY 361 TGGTGGTGTGATGGTGAAGACGAGCGGTGGAAGCGCTGTAAGCGCGGTGCACAATC 420
DB 361 TGGTGGTGTGATGGTGAAGACGAGCGGTGGAAGCGCTGTAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGGTGAGTGGGTGATCAATTAATCTCGCTGGATGACAGAGTGCCA 480
DB 421 TTCTCGCGCAACCGGTGAGTGGGTGATCAATTAATCTCGCTGGATGACAGAGTGCCA 480
QY 481 TTGCTGGAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACGAGA 540
DB 481 TTGCTGGAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACGAGA 540

Db 481 TTGCTGGAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
QY 601 TGGTGGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTCTCGG 660
DB 601 TGGTGGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTCTCGG 660
QY 661 CGCGTCTGCGTCTGCGTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGCGATAG 720
DB 661 CGCGTCTGCGTCTGCGTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGCGATAG 720
QY 721 CGGAACCGGAAGGCGACTGGAGTCCCATGTCGGTTCCTCAACCAACCATGCAATGCTGA 780
DB 721 CGGAACCGGAAGGCGACTGGAGTCCCATGTCGGTTCCTCAACCAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGTGCGCAACGATCAGATGGCGTGGCGCAA 840
DB 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGTGCGCAACGATCAGATGGCGTGGCGCAA 840
QY 841 TGGCGCGCATTTACCGAGTCCGGCTGCGCTGCGGATATCTCGTGGATGATGCGATGCG 900
DB 841 TGGCGCGCATTTACCGAGTCCGGCTGCGCTGCGGATATCTCGTGGATGATGCGATGCG 900
QY 901 ACGATACCGAAGACAGTCAATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGTCAATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACCGAGTCCGGCTGCGCTGCGGATATCTCGTGGATGATGCGATGCG 1020
DB 961 GCCTGCTGGGGCAAAACCGAGTCCGGCTGCGCTGCGGATATCTCGTGGATGATGCGATGCG 1020
QY 1021 AGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAAACAACCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAAACAACCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCGCTCTCCCGCGCTTGGCGGATTCATTAATGCACTGGCAAGCAGGTTT 1140
DB 1081 CGCAACCGCGCTCTCCCGCGCTTGGCGGATTCATTAATGCACTGGCAAGCAGGTTT 1140
QY 1141 CCGCACTGGAAGACGGCGAGTGAAGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
DB 1141 CCGCACTGGAAGACGGCGAGTGAAGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAATTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGAGCCATCGGAAGCTGTTATGGCTGTCGAGTTCGTAATCCTGCATTAATTCG 1320
DB 1261 TCAGGCGAGCCATCGGAAGCTGTTATGGCTGTCGAGTTCGTAATCCTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGCTTCGGATTAATGTTTTCGGCGGACATCAACCGGTT 1380
DB 1321 TGTGCTCAAGCGCACTCCCGCTTCGGATTAATGTTTTCGGCGGACATCAACCGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGAACAATTAATCATCGGCTCGTAAATGCTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGAACAATTAATCATCGGCTCGTAAATGCTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCGCTGCTGTTAGGTGTTTTCACCA 1500
DB 1441 ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCGCTGCTGTTAGGTGTTTTCACCA 1500
QY 1501 GCATTTCAACCAAGGACCATAGATTATGAAGCTGAAGAGGTAAGTGGTAAATCTCG 1560
DB 1501 GCATTTCAACCAAGGACCATAGATTATGAAGCTGAAGAGGTAAGTGGTAAATCTCG 1560
QY 1561 ATTTACGGCGATTAAGGCTATAACCGTCTCGCTGAGTCCGTAAGAAATTCAGAAAGAT 1620
DB 1561 ATTTACGGCGATTAAGGCTATAACCGTCTCGCTGAGTCCGTAAGAAATTCAGAAAGAT 1620

QY 1621 ACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAAGAGAAAATTCACACAGGTT 1680
Db |||||||
QY 1621 ACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAAGAGAAAATTCACACAGGTT 1680
Db |||||||
QY 1681 GCGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db |||||||
QY 1681 GCGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db |||||||
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db |||||||
QY 1801 CCGTTTACCTGGATGCGGTAGCTTTAAACGCGCAAGCTGATTTGCCATCCGATCCCTGTT 1860
Db |||||||
QY 1801 CCGTTTACCTGGATGCGGTAGCTTTAAACGCGCAAGCTGATTTGCCATCCGATCCCTGTT 1860
Db |||||||
QY 1861 GAAGGCTTATCGCTGATTTATTAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGGAA 1920
Db |||||||
QY 1921 GAGATCCCGCGCTGGATAAGAACTGAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC 1980
Db |||||||
QY 1921 GAGATCCCGCGCTGGATAAGAACTGAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC 1980
Db |||||||
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTGACGCGGGTTATGCTTCAAG 2040
Db |||||||
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTGACGCGGGTTATGCTTCAAG 2040
Db |||||||
QY 2041 TATGAAAACGCGCAAGTACGACATTAAGACGTGGCGGTGGATTAAGCTGGCGGAAAGCG 2100
Db |||||||
QY 2041 TATGAAAACGCGCAAGTACGACATTAAGACGTGGCGGTGGATTAAGCTGGCGGAAAGCG 2100
Db |||||||
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAAACAAACATCAATGAATCGACACCGATTAC 2160
Db |||||||
QY 2161 TCCATCGCAAGAGTGCCTTTAATAAGCGCAAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db |||||||
QY 2221 GCATGGTCCACATCGACACGACGAAAGTAAATGATGTAACGGTACTGCGGACCTTC 2280
Db |||||||
QY 2281 AAGGGTCAACCATCAAAACCGTTCTGTCGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db |||||||
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCTCGAAACTATCTGACTGATGAAGTCTG 2400
Db |||||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGTAGCGCTGAAGTCTTACAGGAAGAG 2460
Db |||||||
QY 2461 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGTAGCGCTGAAGTCTTACAGGAAGAG 2460
Db |||||||
QY 2461 TTGGCGAAAGATCCACGTATTTCGCCGCCACCATGGAAGAACGCCAGAAAGGTGAAATCATG 2520
Db |||||||
QY 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGATGATCAACGCC 2580
Db |||||||
QY 2581 GCCAGGGTGTGAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||||||
QY 2641 AACAAACAACAATAAACAATAAACACACCTCGGGATCGAGGAGGATTTTCAGAAATTC 2700
Db |||||||

QY 2701 GGATCCTCT 2709
Db |||||||
QY 2701 GGATCCTCT 2709
Db |||||||
RESULT 8
US-10-343-859-8
; Sequence 8, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; TITLE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 6648
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
; OTHER INFORMATION: pMAL-c2x
US-10-343-859-8
Query Match 81.8%; Score 2701; DB 19; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGGAAACCTTTCGGGTATGGCATGATAGCCCGGAGAGA 60
Db |||||||
QY 1 CCGACACCATCGAATGGTGGAAACCTTTCGGGTATGGCATGATAGCCCGGAGAGA 60
Db |||||||
QY 61 GTCAATTCCAGGTTGGTGAATGTGAAACCAAGTAACTGATATGATGTCGACAGATGCGG 120
Db |||||||
QY 61 GTCAATTCCAGGTTGGTGAATGTGAAACCAAGTAACTGATATGATGTCGACAGATGCGG 120
Db |||||||
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTCTGCGAAA 180
Db |||||||
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTCTGCGAAA 180
Db |||||||
QY 181 CCGGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGC 240
Db |||||||
QY 181 CCGGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGC 240
Db |||||||
QY 241 AACAACTCGCGGCAAAACAGTCTGTTGTTGGGTTGCCACCTCCAGTCTGGCCCTGCG 300
Db |||||||
QY 241 AACAACTCGCGGCAAAACAGTCTGTTGTTGGGTTGCCACCTCCAGTCTGGCCCTGCG 300
Db |||||||
QY 301 ACGGCCGCTCGCAAAATTTGTCGCGCGATTAATCTCGCGCCGATCAACTGGGTGCGCAGG 360
Db |||||||
QY 301 ACGGCCGCTCGCAAAATTTGTCGCGCGATTAATCTCGCGCCGATCAACTGGGTGCGCAGG 360
Db |||||||
QY 361 TGGTGGTTCGATGTTAGAAAGAGCGCGCTGAAAGCCCTGTAAGCGCGGTGCAATTC 420
Db |||||||
QY 361 TGGTGGTTCGATGTTAGAAAGAGCGCGCTGAAAGCCCTGTAAGCGCGGTGCAATTC 420
Db |||||||
QY 421 TTCTCGCGCAACGCGTCAGTGGCTGATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
Db |||||||
QY 421 TTCTCGCGCAACGCGTCAGTGGCTGATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
Db |||||||
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db |||||||
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db |||||||
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCACTGGCGGTGGAGCATC 600
Db |||||||

Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGTTACGCGACTGGCGTGGAGCATC 600
Qy 601 TGGTCGATTTGGTTCACAGCAAAATCGCGCTGTGTAGCGGCGCCATTAAGTTCCTGCTCGG 660
Db 601 TGGTCGATTTGGTTCACAGCAAAATCGCGCTGTGTAGCGGCGCCATTAAGTTCCTGCTCGG 660
Qy 661 CGGCTCTGGCTGTGGCTGGCTGGCTGCAATAATCTCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGGCTGTGGCTGGCTGGCTGCAATAATCTCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGGACTGAGTGGCATGTCGCGTTCGCTGTTTCAAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGGACTGAGTGGCATGTCGCGTTCGCTGTTTCAAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCTGTTCCCATCTGGCATGCTGTTGTCGCAACGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCTGTTCCCATCTGGCATGCTGTTGTCGCAACGATCAGATGGCGCTGGCGCGAA 840
Qy 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGTGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGTGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGCGGCTCTCACTGTTGAAAGAAAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGGCTCTCACTGTTGAAAGAAAAACCAACCTGGCGGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTGCAGAGCTTATCATGAGCTGCACGCTGCACCAATGCTTCGCG 1260
Db 1201 GCACAATTCATGTTTGCAGAGCTTATCATGAGCTGCACGCTGCACCAATGCTTCGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGCTGAGGTGCTAAATCACTGCATAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGCTGAGGTGCTAAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACCTCCCGTTCGGATAATGTTTTCGCGCGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTTCGGATAATGTTTTCGCGCGACATCATACCGGTT 1380
Qy 1381 CTGCAAAATTTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGCTAATGTTGGA 1440
Db 1381 CTGCAAAATTTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGCTAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCATTCACCAACAGGACCATAGTATTGAAATCTGAATGTTTTCGCGCGACATCATACCGGTT 1560
Db 1501 GCATTCACCAACAGGACCATAGTATTGAAATCTGAATGTTTTCGCGCGACATCATACCGGTT 1560
Qy 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAGTCCGTTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAGTCCGTTAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680

Qy 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGSCCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGSCCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCGAAGCTGATTGCTTACC CGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCGAAGCTGATTGCTTACC CGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAAAGAGATCTGCTGCGAAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAAAGAGATCTGCTGCGAAACCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAAACGTTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAAACGTTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAAGCTGCTTTTAAAGCGAAACAGCGATGACCATCAACCGCCCGTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCTTTTAAAGCGAAACAGCGATGACCATCAACCGCCCGTGG 2220
Qy 2221 GCATGTCCAAACATCGACACAGCAAGTGAATTAATGCTGTAACGCTGCTGCGGACTTC 2280
Db 2221 GCATGTCCAAACATCGACACAGCAAGTGAATTAATGCTGTAACGCTGCTGCGGACTTC 2280
Qy 2281 AAGGTCACCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
Db 2281 AAGGTCACCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
Qy 2341 CCGAACAAAGAGCTGCAAAAGAGTTTCTCGAATACTATCTGCTGATGATGAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGCAAAAGAGTTTCTCGAATACTATCTGCTGATGATGAGTCTG 2400
Qy 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTTGGCGAAAGATCCAGTATTGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTTGGCGAAAGATCCAGTATTGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCGTGGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCGTGGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 2581 GCCAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Qy 2641 AACAAACAAACAAATAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 2700
Db 2641 AACAAACAAACAAATAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2700
Qy 2701 GGATCTCT 2709
Db 2701 GGATCTCT 2709

RESULT 9
US-10-343-859-9
; Sequence 9, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCI/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9191
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-muts
; OTHER INFORMATION: plasmid
US-10-343-859-9
Query Match 81.8%; Score 2698; DB 19; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGCATGATAGCGCCCGGAAGAGA 60
DB 1 CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGTGAATGTGAACACGATGAACGTTATACGATGTCGAGAGTATGCGG 120
DB 61 GTCATTCAGGGTGTGAATGTGAACACGATGAACGTTATACGATGTCGAGAGTATGCGG 120
QY 121 GTGTCTCTTATCAGACCGGTTTCCGGGTGTGAACACGAGCCAGCCAGCTTTCTGCGAAAA 180
DB 121 GTGTCTCTTATCAGACCGGTTTCCGGGTGTGAACACGAGCCAGCCAGCTTTCTGCGAAAA 180
QY 181 CCGGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
DB 181 CCGGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAACACAGTGTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAACACAGTGTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTTGTCGGGGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGCGCGTTCGCAAAATTTGTCGGGGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTTAGAAACGAGCGGCTGAAAGCGGCTGTAAGCGGGGTGCACAATC 420
DB 361 TGGTGGTGTGATGTTAGAAACGAGCGGCTGAAAGCGGCTGTAAGCGGGGTGCACAATC 420
QY 421 TTCTCGGCAACCGCTCAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480
DB 421 TTCTCGGCAACCGCTCAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGAAGCTGCCTCACTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
DB 481 TTGCTGTGAAGCTGCCTCACTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGGGCTGCGGCTGCGAGCATC 600
DB 541 CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGGGCTGCGGCTGCGAGCATC 600
QY 601 TGGTGGCATTTGGGTGCACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTCTGTCTCGG 660
DB 601 TGGTGGCATTTGGGTGCACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTCTGTCTCGG 660

DB 601 TGGTGGCATTTGGGTGCACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTCTGTCTCGG 660
QY 661 CCGGTCTCGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CCGGTCTCGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAAGTGCATGTCGGGTTTCAACAAACCATGCAAAATTCGTA 780
DB 721 CGGAACGGGAAGGCGACTGGAAGTGCATGTCGGGTTTCAACAAACCATGCAAAATTCGTA 780
QY 781 ATGAGGCGCATGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGGCA 840
DB 781 ATGAGGCGCATGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGGCA 840
QY 841 TGGCGCCATTACCGAGTCCGGGCTGCGGCTTGGTGGGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGCCATTACCGAGTCCGGGCTGCGGCTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACCCATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACCCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGTGA 1020
DB 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCTCTCACTGCTGTAAGAAAGAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCCCTCTCACTGCTGTAAGAAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGGTTT 1140
DB 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGGTTT 1140
QY 1141 CCGGACTCGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTAGTCTACTCATTTAG 1200
DB 1141 CCGGACTCGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTAGTCTACTCATTTAG 1200
QY 1201 GCACAAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTCTGGCG 1260
DB 1201 GCACAAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTCTGGCG 1260
QY 1261 TCAGGCGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAAATCACTGATTAATTCG 1320
DB 1261 TCAGGCGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAAATCACTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGATTAATGTTTTCGCGCGAGCATCAATACGGTT 1380
DB 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGATTAATGTTTTCGCGCGAGCATCAATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTAGGTGTTTTCACGA 1500
DB 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
DB 1501 GCACCTTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
QY 1561 ATTAACGGCGATTAAGGCTGTAACCGGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATTAAGGCTGTAACCGGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGGTTGAGCATCCGGAATAAATGGAAGAGAAATTCGCCAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGGTTGAGCATCCGGAATAAATGGAAGAGAAATTCGCCAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGCGCTTTGGTGGCTAC 1740

```
QY 1741 GCTCAATCTGGCTGTGGCTGAAATACACCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATACACCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
QY 1801 CGTTTACCTGGATGCGCTGATTAACACGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CGTTTACCTGGATGCGCTGATTAACACGCAAGCTGATGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCGCCAAACCTTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCGCCAAACCTTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGACGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGAGCGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGGCAAGTACGACATTAAGACGCTGGCGTGAATGATGCTGCGGATGATG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTAAGACGCTGGCGTGAATGATGCTGCGGATGATG 2100
QY 2101 GGTCTGACCTTCTGCTGCTGATTAAGAACTGAAGACGAAAGGTAAGAGCGCGCTGATTTCAAC 2160
Db 2101 GGTCTGACCTTCTGCTGCTGATTAAGAACTGAAGACGAAAGGTAAGAGCGCGCTGATTTCAAC 2160
QY 2161 TCCATCGCAGAGCTGCTTAAATGAAGCGGAAACAGCGATGACCATCAACGGCGCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTAAATGAAGCGGAAACAGCGATGACCATCAACGGCGCGTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAAGAACTGAAGACGAAAGGTAAGAGCGCGCTGATTTCAAC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAAGAACTGAAGACGAAAGGTAAGAGCGCGCTGATTTCAAC 2280
QY 2281 AAGGTCACCATCAACACCGTTCGTTGGCGTGTGAGCGAGGATTAACCGCGCCAGT 2340
Db 2281 AAGGTCACCATCAACACCGTTCGTTGGCGTGTGAGCGAGGATTAACCGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAGAGGTTCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAGAGGTTCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGCAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGCAAGAG 2460
QY 2461 TTGGGAAAGATCCACGTTATGCGCGCAACATGGAACCGCCAGAAAGGTTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTTATGCGCGCAACATGGAACCGCCAGAAAGGTTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATCGCGTGAAGACGCGGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATCGCGTGAAGACGCGGATCAACGCC 2580
QY 2581 GCCAGCGTCTGAGACTGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGAGACTGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACACACATACATTAACAACTCGGGATCGAGGAAAGGATTCAGAAATTC 2700
Db 2641 AACCAACACACATACATTAACAACTCGGGATCGAGGAAAGGATTCAGAAATTC 2700
QY 2701 GGATCC 2706
Db 2701 GGATCC 2706
```

RESULT 10
US-10-263-153-40
./ Sequence 40, Application US/10263153

```
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USRS THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 7112
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3189)
; OTHER INFORMATION: pMBP-c2X-Toxop30del11 (52-214aa)
US-10-263-153-40
```

```
Query Match      81.58; Score 2690.4; DB 18; Length 7112;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTCGAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGTCGAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCATTCAGGTTGGTGAATGTGAACACGTTATACGATGTGCGAGATATGCCG 120
Db 61 GTCATTCAGGTTGGTGAATGTGAACACGTTATACGATGTGCGAGATATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCGGTGTGAACAGCGCCAGCGACGTTTCTGCGAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCGGTGTGAACAGCGCCAGCGACGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACATCTGCGCGGCAACACAGTCTGCTGATGCGGCTTGGCACCTCAGTCTGGCCCTGC 300
Db 241 AACATCTGCGCGGCAACACAGTCTGCTGATGCGGCTTGGCACCTCAGTCTGGCCCTGC 300
QY 301 ACGCCGCTCGCAAAATTGTGCGCGCGATTAATTCCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCGCTCGCAAAATTGTGCGCGCGATTAATTCCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGTGTGTGTCGATGTTAGAACGAAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGTGTGTGTCGATGTTAGAACGAAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTGAGTGGCTGATCAATTAATTCCTCGCGCGATCAACTGGGTGCCAG 480
Db 421 TTCTCGCGCAACCGCTGAGTGGCTGATCAATTAATTCCTCGCGCGATCAACTGGGTGCCAG 480
QY 481 TTGCTGTGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACGAGA 540
Db 481 TTGCTGTGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACGAGA 540
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGAGCGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGTGTGCTGCTGCGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTTAAGTCTGTCTCGG 660
Db 601 TGTGTGCTGCTGCGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTTAAGTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
```


Db 661 |||||CGGCTCTGGCTGGCTGGCAATAATCTCCTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGGGATGCTGGTTGCCAAGATCAGATGGCGTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTGGGATGCTGGTTGCCAAGATCAGATGGCGTGGCGGCA 840
Qy 841 TGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
Qy 961 GCCTCTGGGGCAACCAAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Db 961 GCCTCTGGGGCAACCAAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Qy 1021 AGGGCAATCAGTGTGCGCGTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGTGTGCGCGTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGGCAGCACAGGTTT 1140
Db 1081 CGCAACCGCCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGGCAGCACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTCAGAGCTTATCATGACTGCAGGTGCACCAATGCTTCGGG 1260
Db 1201 GCACAATTCATGTTTCAGAGCTTATCATGACTGCAGGTGCACCAATGCTTCGGG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGATAATCTCATTAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGATAATCTCATTAATTCG 1320
Qy 1321 TGTCTCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCAGCATCAATACGGTT 1380
Db 1321 TGTCTCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCAGCATCAATACGGTT 1380
Qy 1381 CTGGCAATATCTGAAATGAGCTGTTGACATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGACATTAATCATCGGCTCGTATAATGTGTGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCAGGA 1500
Db 1441 ATTGTAGCGGATAACAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCAGGA 1500
Qy 1501 GCATTTCAACAAAGGACCATAGATTAATGAAAACCTGAAGAGGTAATCTGTAATCTGG 1560
Db 1501 GCATTTCAACAAAGGACCATAGATTAATGAAAACCTGAAGAGGTAATCTGTAATCTGG 1560
Qy 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATAAAGGCTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGTTGAGATCCGATTAACCTGGAAGAGAAATTCACACAGTTT 1680
Db 1621 ACCGGAATTAAGTCAACGTTGAGATCCGATTAACCTGGAAGAGAAATTCACACAGTTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGCAACCGGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGCAACCGGTTCCAGGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTAACAACGGCAAGCTGATTGTTACCGGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTAACAACGGCAAGCTGATTGTTTACCGGATCGCTGT 1860
Qy 1861 GAAGGTTATCGCTGATTATAACAAAGATCTGCTGCGAACCCGCGCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTATAACAAAGATCTGCTGCGAACCCGCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTGAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTGAAGAGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCTGATCTTCACTGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCTGATCTTCACTGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAANAACGCAAGTACGACATTAAGACGTGGGCGTGGATTAAGCTGGCGGCAAGCG 2100
Db 2041 TATGAANAACGCAAGTACGACATTAAGACGTGGGCGTGGATTAAGCTGGCGGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGTTCCAAATCGACACCAAGTAATTAATGTTGTTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGTTCCAAATCGACACCAAGTAATTAATGTTGTTAAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCCCCAGT 2340
Db 2281 AAGGTTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCCCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAGATCCACGATTTGCGCCCATATGGAAGACCCAGAAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGATTTGCGCCCATATGGAAGACCCAGAAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Qy 2581 GCGAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCGAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Qy 2641 AACACACACAAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTTCAAAATTC 2700
Db 2641 AACACACACAAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTTCAAAATTC 2700

RESULT 11
US-10-263-153-30
; Sequence 30, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.

; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
 ; FILE REFERENCE: 6984.US.01
 ; CURRENT APPLICATION NUMBER: US/10/263,153
 ; CURRENT FILING DATE: 2002-10-22
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 30
 ; LENGTH: 7259
 ; TYPE: DNA
 ; ORGANISM: Toxoplasma gondii
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1528)...(3336)
 ; OTHER INFORMATION: pMBP-c2X-Toxop30del14del18 (83-294aa)
 US-10-263-153-30

Query Match 81.5%; Score 2690.4; DB 18; Length 7259;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGCACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
DB	1	CGCACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGTGAATGTGAACCAAGTAAGCTTATACGATGTGCGAGAGTATGCCG	120
DB	61	GTCAATTCAGGGTGTGAATGTGAACCAAGTAAGCTTATACGATGTGCGAGAGTATGCCG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCGGCGAGCCAGCCAGCTTTCTCGGAAA	180
DB	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCGGCGAGCCAGCCAGCTTTCTCGGAAA	180
QY	181	CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCGMAACAGTCGTTGTGATGCGGTTGCGACCTCCAGTCTGCGCCCTGC	300
DB	241	AACAACTGGCGGCGMAACAGTCGTTGTGATGCGGTTGCGACCTCCAGTCTGCGCCCTGC	300
QY	301	ACGCGCGTCCGCAATTTGTCGCGCGATTAATCTCCGCGCGATCACTGGGTGCGACGC	360
DB	301	ACGCGCGTCCGCAATTTGTCGCGCGATTAATCTCCGCGCGATCACTGGGTGCGACGC	360
QY	361	TGGTGGTTCGATGGTAGAACGAAGCGCGTGAAGCTGTGAAGCGCGGTGCACAATC	420
DB	361	TGGTGGTTCGATGGTAGAACGAAGCGCGTGAAGCTGTGAAGCGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCGAGTGCCA	480
DB	421	TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCGAGTGCCA	480
QY	481	TTGCTGTGGAAGCTCCGTCACATTTGTCGCGGTTATTTCTTGATGTCTTGACCGAGA	540
DB	481	TTGCTGTGGAAGCTCCGTCACATTTGTCGCGGTTATTTCTTGATGTCTTGACCGAGA	540
QY	541	CACCCATCAACAGTATTATTTTCCCAATGAAGACGCTACCGAGCTGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTATTTTCCCAATGAAGACGCTACCGAGCTGGGCGTGGAGCATC	600
QY	601	TGGTGGCATTTGGGTACACAGCAATCCGCTGTGTAGCGGCGCATTAAGTCTCTCGG	660
DB	601	TGGTGGCATTTGGGTACACAGCAATCCGCTGTGTAGCGGCGCATTAAGTCTCTCGG	660
QY	661	CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGACTGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780
DB	721	CGGAACGGGAAGCGACTGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780

QY	781	ATGAGGCATCGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCGTGGCGCGAA	840
DB	781	ATGAGGCATCGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCGTGGCGCGAA	840
QY	841	TGCGCGCATTTACCGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGCGCATTTACCGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACACCATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACACCATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAAAACAGCGTGTGCAACTCTCTCAGGCGCAGCGGTGA	1020
DB	961	GCCTGCTGGGCAAAACAGCGTGTGCAACTCTCTCAGGCGCAGCGGTGA	1020
QY	1021	AGGCAATACAGTGTGTCCTCTCACTGCTGAAAAAGAAAAACACCTCGCGCCCAATA	1080
DB	1021	AGGCAATACAGTGTGTCCTCTCACTGCTGAAAAAGAAAAACACCTCGCGCCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAAGACAGTTT	1140
DB	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAAGACAGTTT	1140
QY	1141	CCGACTTGAAGCGGCGAGTGAAGCAACGCAATTAATGAGTTAGCTCCTCATTTAG	1200
DB	1141	CCGACTTGAAGCGGCGAGTGAAGCAACGCAATTAATGAGTTAGCTCCTCATTTAG	1200
QY	1201	GCAAAATTCATGTTGACAGCTTATCATCGACTGCAAGGTGCACCAATTCCTTGGCG	1260
DB	1201	GCAAAATTCATGTTGACAGCTTATCATCGACTGCAAGGTGCACCAATTCCTTGGCG	1260
QY	1261	TCAGGAGCCATCGGAAGCTGTGATGGCTGTGCGAGTGTGCACTGCACTGCAATTCG	1320
DB	1261	TCAGGAGCCATCGGAAGCTGTGATGGCTGTGCGAGTGTGCACTGCACTGCAATTCG	1320
QY	1321	TGTCGCTCAAGCGGCACTCCCGTCTTGGAATAATGTTTTTGGCGCGACATCAACGGTT	1380
DB	1321	TGTCGCTCAAGCGGCACTCCCGTCTTGGAATAATGTTTTTGGCGCGACATCAACGGTT	1380
QY	1381	CTGGCAAAATTCGAAATGAGCTGTGTAATTAATCATCGCTCGTAAATGTTGTTGA	1440
DB	1381	CTGGCAAAATTCGAAATGAGCTGTGTAATTAATCATCGCTCGTAAATGTTGTTGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTTACACAGGAACAGCCAGTCCGTTTAGTGTTTTTCA	1500
DB	1441	ATTGTGAGCGGATAACAATTTTACACAGGAACAGCCAGTCCGTTTAGTGTTTTTCA	1500
QY	1501	GCACTTCAACCAAGGACCATAGCATATGAAATCGAAGAGGTAAACTGGTAAATCTCG	1560
DB	1501	GCACTTCAACCAAGGACCATAGCATATGAAATCGAAGAGGTAAACTGGTAAATCTCG	1560
QY	1561	ATTACCGGATAAAGGCTTAAACCGTCTCGCTGAGTCCGTTAGTAAAGAAATTCGA	1620
DB	1561	ATTACCGGATAAAGGCTTAAACCGTCTCGCTGAGTCCGTTAGTAAAGAAATTCGA	1620
QY	1621	ACCGAATTTAAAGTCAACCGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACCA	1680
DB	1621	ACCGAATTTAAAGTCAACCGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACCA	1680
QY	1681	GGGCAATTCGGGATGCGCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGGCTAC	1740
DB	1681	GGGCAATTCGGGATGCGCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAATACCCCGGCAAGCGTTCGAGGACAGCTGTAT	1800
DB	1741	GCTCAATCTGGCTGTGGCTGAATACCCCGGCAAGCGTTCGAGGACAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATCGGATACCAACCGCAAGCTGATTTACCCGATCGCTGT	1860
DB	1801	CGGTTTACCTGGGATCGGATACCAACCGCAAGCTGATTTACCCGATCGCTGT	1860
QY	1861	GAAGCGTTATCGCTGATTTATTAACAAAGATCTGCTGCCGAACCCGCCAAAAA	1920

```
Db 1861 GAAGGTTATCGCTGATTTATTAACAAGATCGTCGCCGAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGAGCGAAAGGTGAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGAGCGAAAGGTGAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAAGCGTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAAGCGTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAAAAACAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAAAAACAACACATGAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTAAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTAAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTAAACGTTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTAAACGTTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCAATCCAAACCGTTCGTCGCGTGGTGAAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTCGTCGCGTGGTGAAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGCGTGGTGAAGCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGCGTGGTGAAGCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCACGATTTCGCCCCACCTGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTCGCCCCACCTGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAATATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
Db 2521 CCGAATATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
```

RESULT 12

```
US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq For Windows Version 4.0
```

```
; SEQ ID NO 35
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528) ... (3399)
; OTHER INFORMATION: pMBP-c2X-Toxop30del10 (52-284aa)
US-10-263-153-35
```

```
Query Match
Best Local Similarity 81.5%; Score 2690.4; DB 18; Length 7322;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGTAGGCCCGGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGTAGGCCCGGGAAGAGA 60
Qy 61 GTCAAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTCGCAGAGTATGCCG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTCGCAGAGTATGCCG 120
Qy 121 GTGTCCTTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCACGTTTCTGCGAAAA 180
Db 121 GTGTCCTTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCACGTTTCTGCGAAAA 180
Qy 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGCCCTGCG 300
Db 241 AACAACTGGCGGGCAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGCCCTGCG 300
Qy 301 ACGCGCGTTCGCAAAATTCGTCGCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTCGTCGCGCGGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGGTTCGATGATAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTTCGATGATAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACGGGTGATGGGTGATCATTAATCTCCGTGATGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACGGGTGATGGGTGATCATTAATCTCCGTGATGATGACCAAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGCGTATTCTTCTGATGTTCTGACCA 540
Db 481 TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGCGTATTCTTCTGATGTTCTGACCA 540
Qy 541 CACCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGGTGGAGCATC 600
Db 541 CACCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGGTACACAGCAAAATTCGGCTGTTAGCGGGGCCATTAAAGTTCTGTCTGG 660
Db 601 TGGTGCATTTGGGTACACAGCAAAATTCGGCTGTTAGCGGGGCCATTAAAGTTCTGTCTGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGAAACGGGAAGGGACCTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGAAACGGGAAGGGACCTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCTCCACTCGCATGCTGGTTGCAACAGATCAGATGCGGTGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCTCCACTCGCATGCTGGTTGCAACAGATCAGATGCGGTGGCGCAA 840
Qy 841 TGGCGCCATTAACCGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATAGC 900
Db 841 TGGCGCCATTAACCGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATAGC 900
```

Qy	901	ACGATACCGAAGACACAGCTCATGTTATATATCCGCGCTTAAACACCACTCAAAACAGGATTTTC	960
Db	901	ACGATACCGAAGACACAGCTCATGTTATATCCGCGCTTAAACACCACTCAAAACAGGATTTTC	960
Qy	961	GCCTGCTGGGCAAAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020
Db	961	GCCTGCTGGGCAAAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTTGCGCGCTCTCACTGTGTAAAGAAACACACCTCGCGGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCGCGCTCTCACTGTGTAAAGAAACACACCTCGCGGCCCAATA	1080
Qy	1081	CGCAAAACCGCTCTCCCGCGCGTGTGGCCGATTCAATTAATGCAGCTGGCAGCACAGGTTT	1140
Db	1081	CGCAAAACCGCTCTCCCGCGCGTGTGGCCGATTCAATTAATGCAGCTGGCAGCACAGGTTT	1140
Qy	1141	CCCGACTGGAAAGCGGCGAGTGAGCGGCAACGCAATTAATGTGAGTTAGTCACTCATTAG	1200
Db	1141	CCCGACTGGAAAGCGGCGAGTGAGCGGCAACGCAATTAATGTGAGTTAGTCACTCATTAG	1200
Qy	1201	GCACAATTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
Db	1201	GCACAATTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
Qy	1261	TCAGGCAGCCATCGAAGCTGTGTATGTGCTGTGCAGGTCGTAATCACTGCATAATTGCG	1320
Db	1261	TCAGGCAGCCATCGAAGCTGTGTATGTGCTGTGCAGGTCGTAATCACTGCATAATTGCG	1320
Qy	1321	TGTCGCTCAAGGCGCACTCCCGTTCTCGATAATGTTTTTGGCGCGACATCAACCGTT	1380
Db	1321	TGTCGCTCAAGGCGCACTCCCGTTCTCGATAATGTTTTTGGCGCGACATCAACCGTT	1380
Qy	1381	CTGCGCAATATTTCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
Db	1381	CTGCGCAATATTTCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
Qy	1441	ATTGTGAGCGGATAACAATTTTCACACAGGAAACAGCCAGTCGTTTGGTGTGTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAATTTTCACACAGGAAACAGCCAGTCGTTTGGTGTGTTTTCACGA	1500
Qy	1501	GCATTTCCACAAACAGGACCATAGATTATGAAACTGAAGAAGTAACTGGTAAATCTCG	1560
Db	1501	GCATTTCCACAAACAGGACCATAGCATATGAAATCGAAGAAGTAACTGGTAAATCTCG	1560
Qy	1561	ATTAAACGGCATAAAGGCTATAAGGCTCTCGCTGAAAGTGGTAAGAAATTCGAGAAGAT	1620
Db	1561	ATTAAACGGCATAAAGGCTATAAGGCTCTCGCTGAAAGTGGTAAGAAATTCGAGAAGAT	1620
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATCCCAAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATCCCAAGGTT	1680
Qy	1681	CGCGCAACTGGCGATGGCCCTGCATTATCTCTGGGCACACGACCGCTTTGGTGCTTAC	1740
Db	1681	CGCGCAACTGGCGATGGCCCTGCATTATCTCTGGGCACACGACCGCTTTGGTGCTTAC	1740
Qy	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGACAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCCGTACGTTTAAACCGGCAAGCTGATTGCTTACCAGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCCGTACGTTTAAACCGGCAAGCTGATTGCTTACCAGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCAAAAACCTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCAAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980

Qy	1981	CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACGGGGGTATTGCGTTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACGGGGGTATTGCGTTTCAAG	2040
Qy	2041	TATGAAAAACGCAAGTAGACACATTTAAAGACGTTGGGGGTGGATTAACGTTGGCGCGAAGCG	2100
Db	2041	TATGAAAAACGCAAGTAGACACATTTAAAGACGTTGGGGGTGGATTAACGTTGGCGCGAAGCG	2100
Qy	2101	GGTCTGACCTTTCCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGGATTAC	2160
Db	2101	GGTCTGACCTTTCCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGGATTAC	2160
Qy	2161	TCCATCCGAGAGCTGCCTTTTAATAAGGCGAACACAGCGATGACCATCAACGGCCCGGTGG	2220
Db	2161	TCCATCCGAGAGCTGCCTTTTAATAAGGCGAACACAGCGATGACCATCAACGGCCCGGTGG	2220
Qy	2221	GCATGGTCCAAACATCGACACACAGCAAAAGTGAATTAATGGTGTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGGTCCAAACATCGACACACAGCAAAAGTGAATTAATGGTGTAAACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGGTCAACCATCCAAACCGTTTCGTTGGGGTCTGAGCGCAGGTATTAAACGGCCGAGT	2340
Db	2281	AAGGGTCAACCATCCAAACCGTTTCGTTGGGGTCTGAGCGCAGGTATTAAACGGCCGAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTTCCTCGAAAACTATCTGCTGACTCATGAAGTCTG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTTCCTCGAAAACTATCTGCTGACTCATGAAGTCTG	2400
Qy	2401	GAAGCGGTTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG	2460
Db	2401	GAAGCGGTTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG	2460
Qy	2461	TTGGCGAAAGATCCACGTTATTCGCGCCACCATGGAAAAACGCCAGAAAGGTCAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTTATTCGCGCCACCATGGAAAAACGCCAGAAAGGTCAATCATG	2520
Qy	2521	CCGAAATATCCGCGAGATGTCGGTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAAATATCCGCGAGATGTCGGTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGTCAGACTGTCATGAAGCCCTGAAAGACGCCGACACTAATTTTCGAGCTCG	2640
Db	2581	GCCAGCGGTGTCAGACTGTCATGAAGCCCTGAAAGACGCCGACACTAATTTTCGAGCTCG	2640
Qy	2641	AACAAACAAACATTAACATTAACAAACCACTTCGGGATCGAGGGGAAGGATTTTCAGAATTC	2700
Db	2641	AACAAACAAACATTAACATTAACAAACCACTTCGGGATCGAGGGGAAGGATTTTCAGAATTC	2700

```

RESULT 13
US-10-263-153-25
; Sequence 25, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 7352
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528) ... (3429)

```

; OTHER INFORMATION: pMBP-c2X-ToxoP30del14C (52-294aa)
US-10-263-153-25

Query Match 81.5%; Score 2690.4; DB 18; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGGACCAATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATGAGCCCGGGAAGAGA	60
DB	1	CGGACCAATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATGAGCCCGGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTATACGATCTCCAGAGTATGCG	120
DB	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTATACGATCTCCAGAGTATGCG	120
QY	121	GTGCTCTTATCAGACCCGTTTCCCGCGTGGTGAACCAAGCCAGCCACGCTTCTCGGAAAA	180
DB	121	GTGCTCTTATCAGACCCGTTTCCCGCGTGGTGAACCAAGCCAGCCACGCTTCTCGGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTTACATTCCTCCAAACCGCGTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTTACATTCCTCCAAACCGCGTGGCAC	240
QY	241	AACAACTGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCCACTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCCACTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCCGTCGCAAAATTTGCGCGCGATTAATCTCGCGCGCATCAACTGGGTGCGACGG	360
DB	301	ACGCGCCGTCGCAAAATTTGCGCGCGATTAATCTCGCGCGCATCAACTGGGTGCGACGG	360
QY	361	TGGTGGTTCGATGTAGAACGAAAGCGGCTCGAAGCTGTAAAGCGGCGTGCACAATC	420
DB	361	TGGTGGTTCGATGTAGAACGAAAGCGGCTCGAAGCTGTAAAGCGGCGTGCACAATC	420
QY	421	TTCTCGCCCAACGCTCAGTGGCTGATCAATTAATCTCCGCTGATGACCAAGGATGCCA	480
DB	421	TTCTCGCCCAACGCTCAGTGGCTGATCAATTAATCTCCGCTGATGACCAAGGATGCCA	480
QY	481	TTGCTGTGGAACTGCCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTGACCCAGA	540
DB	481	TTGCTGTGGAACTGCCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACCGACTGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACCGACTGGGCGTGGAGCATC	600
QY	601	TGGTGGCATTTGGGTACCAAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTGG	660
DB	601	TGGTGGCATTTGGGTACCAAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTGG	660
QY	661	CGCGTCTGGTCTGGTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGGTCTGGTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAAACGGAAGCGCATGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATGCTGA	780
DB	721	CGGAAACGGAAGCGCATGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTTCACCATGCGATGCTGGTGGCCAAACGATCAGATGGGCTGGGGGCA	840
DB	781	ATGAGGGCATCGTTCACCATGCGATGCTGGTGGCCAAACGATCAGATGGGCTGGGGGCA	840
QY	841	TGCGGCCATTAACCGAGTCCGGCTGCGGTTGGTGGATATCTCGGTAGTGGGATAG	900
DB	841	TGCGGCCATTAACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACACGAGTGGACCGCTTGGTGCACACTCTCTCAGGCGCCAGGCGGTGA	1020

DB	961	GCCTGCTGGGGCAACACGAGCGTGAGCCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020
QY	1021	AGGCAATACAGCTGTGTCCCGTCTCAGCTGTGTAAGAAAAACCACTCTGCGGCCAATA	1080
DB	1021	AGGCAATACAGCTGTGTCCCGTCTCAGCTGTGTAAGAAAAACCACTCTGCGGCCAATA	1080
QY	1081	CGAAAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATCAGCTGCGACGACGAGTTT	1140
DB	1081	CGAAAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATCAGCTGCGACGACGAGTTT	1140
QY	1141	CCGACTTGAAGAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
DB	1141	CCGACTTGAAGAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATCTCTCTGCGG	1260
DB	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATCTCTCTGCGG	1260
QY	1261	TCAGGACCCATCGGAAGCTGTGTATGGCTGTGAGGTCTGAATCACTGCATTAATTCG	1320
DB	1261	TCAGGACCCATCGGAAGCTGTGTATGGCTGTGAGGTCTGAATCACTGCATTAATTCG	1320
QY	1321	TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACCGGTT	1380
DB	1321	TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACCGGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA	1440
DB	1381	CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA	1440
QY	1441	ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCAGTCCGTTTAGGTGTTTTCACGA	1500
DB	1441	ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCAGTCCGTTTAGGTGTTTTCACGA	1500
QY	1501	GCATTTACCAACAAAGGACCATAGATTAATGAAATCTGAAGAAAGGTAAATCTGTAATCTGG	1560
DB	1501	GCATTTACCAACAAAGGACCATAGATTAATGAAATCTGAAGAAAGGTAAATCTGTAATCTGG	1560
QY	1561	ATTAACGCGGATTAAGGCTTAACCGGCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAACGCGGATTAAGGCTTAACCGGCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT	1680
DB	1621	ACCGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT	1680
QY	1681	CGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
DB	1681	CGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATTCGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
DB	1741	GCTCAATTCGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
QY	1801	CGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCCGATCCCTGTT	1860
DB	1801	CGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCCGATCCCTGTT	1860
QY	1861	GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920
DB	1861	GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGCGGGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGCGGGGTTATCGGTTCAAG	2040
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGGTGATTAACGCTGGCGGAAAGCG	2100
DB	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGGTGATTAACGCTGGCGGAAAGCG	2100

QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACACATGAATGCAGACCCGATTAC 2160
DB |||||
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACACATGAATGCAGACCCGATTAC 2160
DB |||||
QY 2161 TCCATGCGCAGAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGCGCGGTGG 2220
DB |||||
QY 2161 TCCATGCGCAGAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGCGCGGTGG 2220
DB |||||
QY 2221 GCATGGTCCAAACATCGACACCGCAAGTGAATTATGGTGAACGCTACTGCGGACCTTC 2280
DB |||||
QY 2221 GCATGGTCCAAACATCGACACCGCAAGTGAATTATGGTGAACGCTACTGCGGACCTTC 2280
DB |||||
QY 2281 AAGGGTCAACCATCCAAACCGTTGCTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
DB |||||
QY 2281 AAGGGTCAACCATCCAAACCGTTGCTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
DB |||||
QY 2341 CGAACAAAGAGCTGCGCAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
DB |||||
QY 2341 CGAACAAAGAGCTGCGCAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
DB |||||
QY 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
DB |||||
QY 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
DB |||||
QY 2461 TTGGCGAAGATCCAGTATTCGCCGCCACCATGGAAGAAACGCCAGAAAGTGAAATCATG 2520
DB |||||
QY 2461 TTGGCGAAGATCCAGTATTCGCCGCCACCATGGAAGAAACGCCAGAAAGTGAAATCATG 2520
DB |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGTTGATGCGGTGCGTACTGCGGTGATCAAGCC 2580
DB |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGTTGATGCGGTGCGTACTGCGGTGATCAAGCC 2580
DB |||||
QY 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGTCTG 2640
DB |||||
QY 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGTCTG 2640
DB |||||
QY 2641 AACAAACAACAATAACAATAACAACACCTCGGATCGAGGGAAGATTCAGAAATTC 2700
DB |||||
QY 2641 AACAAACAACAATAACAATAACAACACCTCGGATCGAGGGAAGATTCAGAAATTC 2700
DB |||||

RESULT 14

US-10-263-153-20
; Sequence 20, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxo30del13C (52-300aa)
US-10-263-153-20

Query Match

Best Local Similarity 81.5%; Score 2690.4; DB 19; Length 7370;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB |||||
QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB |||||
QY 61 GTCAAATTCAGAGTGGTGAATGTAACAGTAACAGTTATACGATGTCGAGAGTATGCGG 120
DB |||||
QY 61 GTCAAATTCAGAGTGGTGAATGTAACAGTAACAGTTATACGATGTCGAGAGTATGCGG 120
DB |||||
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACAGGCGAGCCAGTTCTCGGAAA 180
DB |||||
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACAGGCGAGCCAGTTCTCGGAAA 180
DB |||||
QY 181 CGCGGGAAGAGTGAAGCGCGGCGATGCGGAGCTGAATTAACATTCACCAACCGGTGGCAC 240
DB |||||
QY 181 CGCGGGAAGAGTGAAGCGCGGCGATGCGGAGCTGAATTAACATTCACCAACCGGTGGCAC 240
DB |||||
QY 241 AACAACTGGCGGGGCAACAGTCTGCTGATTCGGCGTTCGCGTTCAGCTCCAGTCTGCGCCCTGC 300
DB |||||
QY 241 AACAACTGGCGGGGCAACAGTCTGCTGATTCGGCGTTCGCGTTCAGCTCCAGTCTGCGCCCTGC 300
DB |||||
QY 301 ACGCCCGTTCGCAAAATGTCGCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB |||||
QY 301 ACGCCCGTTCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
DB |||||
QY 361 TGGTGGTGTGATGAGAGCGGCGTGAAGCGCTGTAAGCGCGGTGCACAATC 420
DB |||||
QY 361 TGGTGGTGTGATGAGAGCGGCGTGAAGCGCTGTAAGCGCGGTGCACAATC 420
DB |||||
QY 421 TTCTCGCGCAACCGTCAAGTGGCTGATCAATTAATCTCGCGTGGATGACAGGATGCCA 480
DB |||||
QY 421 TTCTCGCGCAACCGTCAAGTGGCTGATCAATTAATCTCGCGTGGATGACAGGATGCCA 480
DB |||||
QY 481 TTCTCGTGGAGCTGCTGCACTAATTTTCGGCGCTTATTTCTTGATGTCTCTGACCAGA 540
DB |||||
QY 481 TTCTCGTGGAGCTGCTGCACTAATTTTCGGCGCTTATTTCTTGATGTCTCTGACCAGA 540
DB |||||
QY 541 CACCCATCAACAGTATTTCTCCCATGAAGAGCGTACGCGCTGGCGGCTGGAGCATC 600
DB |||||
QY 541 CACCCATCAACAGTATTTCTCCCATGAAGAGCGTACGCGCTGGCGGCTGGAGCATC 600
DB |||||
QY 601 TGGTGCATTTGGTCAACAGCAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
DB |||||
QY 601 TGGTGCATTTGGTCAACAGCAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
DB |||||
QY 661 CGGCTCTGCTGCTGCGTGGCTGGCATTAATATCTCACTCGCAATCAATTCAGCGCATAG 720
DB |||||
QY 661 CGGCTCTGCTGCTGCGTGGCTGGCATTAATATCTCACTCGCAATCAATTCAGCGCATAG 720
DB |||||
QY 721 CGGAACGGGAAGCGGACGTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
DB |||||
QY 721 CGGAACGGGAAGCGGACGTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
DB |||||
QY 781 ATGAGGCGATCGTTCCTCCATCGGATGCTGGTTGCCAAGATGCGGTGGCGCGCAA 840
DB |||||
QY 781 ATGAGGCGATCGTTCCTCCATCGGATGCTGGTTGCCAAGATGCGGTGGCGCGCAA 840
DB |||||
QY 841 TGGCGGCCATTACCGAGTCCGGCTGCGGTGGTGGCGATATCTCGGTAGTGGATAG 900
DB |||||
QY 841 TGGCGGCCATTACCGAGTCCGGCTGCGGTGGTGGCGATATCTCGGTAGTGGATAG 900
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960
DB |||||
QY 961 GCGTCTGGGCGAAACAGCGTGGAGCGCTTGTGCAACTCTCTCAGGCGCGCGGTGA 1020
DB |||||
QY 961 GCGTCTGGGCGAAACAGCGTGGAGCGCTTGTGCAACTCTCTCAGGCGCGCGGTGA 1020
DB |||||
QY 1021 AGGGCAATCAGTGTGCGCTCTCACTGGTGAAGAAACCAACCCCTGCGCGCCCAATA 1080
DB |||||
QY 1021 AGGGCAATCAGTGTGCGCTCTCACTGGTGAAGAAACCAACCCCTGCGCGCCCAATA 1080
DB |||||
QY 1081 CGCAACCGCCCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGCGACGACAGGTTT 1140
DB |||||

Db 1081 |||||CGCAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGTGGCAGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTTAG 1200
Qy 1201 GCACAATTTCTGATGTTTGACAGCTTATCATGACTGCGAGTGCACCAATGCTTCTGCGG 1260
Db 1201 GCACAATTTCTGATGTTTGACAGCTTATCATGACTGCGAGTGCACCAATGCTTCTGCGG 1260
Qy 1261 TCAGGCGCATCGGAAGCTGTTGATGCTGTGCGAGTGCATTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGCGCATCGGAAGCTGTTGATGCTGTGCGAGTGCATTAATCACTGCATTAATTCG 1320
Qy 1321 TGTCTCTCAAGCGGCACCTCCGCTTCGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTCTCTCAAGCGGCACCTCCGCTTCGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATACAAATTCACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATACAAATTCACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCACCTTCACCAACAGGACCATAGATTTATGAAAACTGAAGAGGTAAACTGGTATCTGG 1560
Db 1501 GCACCTTCACCAACAGGACCATAGATTTATGAAAACTGAAGAGGTAAACTGGTATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGACACAGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGACACAGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGCTGAAATCAACCGGCAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTGCTGAAATCAACCGGCAAGCGTTCCAGGACAAAGCTGAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTACGTTTCAACCGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTCAACCGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGACCTGATTTAAAAACAAACATGATGACAGACACCGGATAC 2160
Db 2101 GGTCTGACCTTCTGCTGACCTGATTTAAAAACAAACATGATGACAGACACCGGATAC 2160
Qy 2161 TCCATCGCAAGAGCTGCCTTTTAAAGCGGAAAACAGCGATACCATCAACCGCCCGTGG 2220

Db 2161 TCCATCGCAAGAGCTGCCTTTTAAAGGCGAAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACCAAGTGAATTAATGGTGAACGGTATCTCCGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACCAAGTGAATTAATGGTGAACGGTATCTCCGACCTTC 2280
Qy 2281 AAGGTCNAACCATCAAAACCGTTTCTGTCGCTGAGCGCAGGTATTAACCGCCCGAGT 2340
Db 2281 AAGGTCNAACCATCAAAACCGTTTCTGTCGCTGAGCGCAGGTATTAACCGCCCGAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAAGCGGTTAATAAGACAAACCCGTCGGTCCGCTAGCGTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAAGCGGTTAATAAGACAAACCCGTCGGTCCGCTAGCGTGAAGTCTTTACGAGGAAG 2460
Qy 2461 TTGGGAAAAGATCCACGTATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAAGATCCACGTATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGGAAGACGCGCAGACTTAATTCGAGCTG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGGAAGACGCGCAGACTTAATTCGAGCTG 2640
Qy 2641 AACAAACAACAATTAACAATAACAACCTCGGGATCGAGGAGGATTTAGAAATTC 2700
Db 2641 AACAAACAACAATTAACAATAACAACCTCGGGATCGAGGAGGATTTAGAAATTC 2700

RESULT 15

US-10-263-153-61
; Sequence 61, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxop30NIX1
US-10-263-153-61

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATCGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATCGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTCAGGTTGTTGATGTGAACCACTTAACCTTATACGATCTCCGAGGATATCGG 120
Db 61 GTCAATTCAGGTTGTTGATGTGAACCACTTAACCTTATACGATCTCCGAGGATATCGG 120

QY	121	GTGCTCTTATCAGACCGGTTTCCGCGTGGTGAACAGGCCAGCCACGTTTCTGCGAAA	180
Db	121	GTGCTCTTATCAGACCGGTTTCCGCGTGGTGAACAGGCCAGCCACGTTTCTGCGAAA	180
QY	181	CGCGGGAAGAGTGGAGCGCGGATCGCGGAGCTGAATTAATTCACACCGCGTGGCAC	240
Db	181	CGCGGGAAGAGTGGAGCGCGGATCGCGGAGCTGAATTAATTCACACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC	300
QY	301	AGCGCGCTGCAGAAATGTTCGCGCGGATTAATCTCGCGCGGATCAATCGGGTGCACGG	360
Db	301	AGCGCGCTGCAGAAATGTTCGCGCGGATTAATCTCGCGCGGATCAATCGGGTGCACGG	360
QY	361	TGCGTGTGTCGATGGTGAACGAGCGCGCTGAAAGCCTGTAAAGCGCGGTGCAAAATC	420
Db	361	TGCGTGTGTCGATGGTGAACGAGCGCGCTGAAAGCCTGTAAAGCGCGGTGCAAAATC	420
QY	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCGCGCGGATCAATCGGGTGC	480
Db	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCGCGCGGATCAATCGGGTGC	480
QY	481	TTGCTGTGGAAGCTGCTGCAATAATGTTTCGGCGTTATTTCTTGATGTTCTCTGACCA	540
Db	481	TTGCTGTGGAAGCTGCTGCAATAATGTTTCGGCGTTATTTCTTGATGTTCTCTGACCA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGGACTGGCGGTGAGCATC	600
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGGACTGGCGGTGAGCATC	600
QY	601	TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCAATTAAGTTCTGTCGG	660
Db	601	TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCAATTAAGTTCTGTCGG	660
QY	661	CGCGTCTGGCTGGCTGGGCAATAATCTCACTCGCAATCAAAATCAGCCGATAG	720
Db	661	CGCGTCTGGCTGGCTGGGCAATAATCTCACTCGCAATCAAAATCAGCCGATAG	720
QY	721	CGAAACGGGAAGGACATGGAAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGAAACGGGAAGGACATGGAAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGGCATCTGTTCCCATCGGATGCTGTTGCAACAGATCAGATGGCGTGGCGCAA	840
Db	781	ATGAGGGCATCTGTTCCCATCGGATGCTGTTGCAACAGATCAGATGGCGTGGCGCAA	840
QY	841	TGGCGCCATTTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGGCGCCATTTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCCATCAAAACAGGATTTTC	960
Db	901	ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCCATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACAGCGTGACCGCTTCTGGAATCTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAACAGCGTGACCGCTTCTGGAATCTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGCAATACAGCTGTTGCCGCTCTCACTGTGTAAGAAAACACCCCTGGCGCCCAATA	1080
Db	1021	AGGCAATACAGCTGTTGCCGCTCTCACTGTGTAAGAAAACACCCCTGGCGCCCAATA	1080
QY	1081	CGCAACCGGCTCTCCCGCGGTTGGCGATTAATTAATGCACTGGCAACGAGGTTT	1140
Db	1081	CGCAACCGGCTCTCCCGCGGTTGGCGATTAATTAATGCACTGGCAACGAGGTTT	1140
QY	1141	CCGACTTGAAGAGCGGCGAGTGAACGCAATTAATGTTAGTTAGTCACTCATTTAG	1200
Db	1141	CCGACTTGAAGAGCGGCGAGTGAACGCAATTAATGTTAGTTAGTCACTCATTTAG	1200

QY	1201	GCACAAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATCTTCTGGCG	1260
Db	1201	GCACAAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATCTTCTGGCG	1260
QY	1261	TCAGCGAGCCATTCGGAAGCTGTGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCG	1320
Db	1261	TCAGCGAGCCATTCGGAAGCTGTGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCCGTTCTCGATAATGTTTTTTCGCGCGACATCATAAACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCCGTTCTCGATAATGTTTTTTCGCGCGACATCATAAACGGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTTGTGACAAATTAATCATCTGGCTCGTAAATGTGTGA	1440
Db	1381	CTGCAAAATTTCTGAAATGAGCTTGTGACAAATTAATCATCTGGCTCGTAAATGTGTGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA	1500
QY	1501	GCATTCACCAACAGGACCATAGATTTATGAAAATCTGAAGAGGTAACTGGTAATCTGG	1560
Db	1501	GCATTCACCAACAGGACCATAGATTTATGAAAATCTGAAGAGGTAACTGGTAATCTGG	1560
QY	1561	ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGGAAGAAT	1620
Db	1561	ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGGAAGAAT	1620
QY	1621	ACCGGAATTTAAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT	1680
Db	1621	ACCGGAATTTAAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT	1680
QY	1681	GCAGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGGTGCTAC	1740
Db	1681	GCAGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGGTGCTAC	1740
QY	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCGGACAAAGCGTTCCAGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCGGACAAAGCGTTCCAGACAGCTGTAT	1800
QY	1801	CGTCTTACTGGGATGCGGTACGTTTAAACGCAAGCTGATTTACCCGATCGCTGTT	1860
Db	1801	CGTCTTACTGGGATGCGGTACGTTTAAACGCAAGCTGATTTACCCGATCGCTGTT	1860
QY	1861	GAAAGCTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAAAGCTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCCAAAACCTGGGAA	1920
QY	1921	GAGATCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC	1980
QY	1981	CTGCAAGAACCTTACCTGGCGCTGATGCTGTGACGGGGTTATCGGTTCAAG	2040
Db	1981	CTGCAAGAACCTTACCTGGCGCTGATGCTGTGACGGGGTTATCGGTTCAAG	2040
QY	2041	TATGAAAACGGCAAGTACGACATTTAAAGAGCTGGCGTGAATACGCTGGCGCGAAGCG	2100
Db	2041	TATGAAAACGGCAAGTACGACATTTAAAGAGCTGGCGTGAATACGCTGGCGCGAAGCG	2100
QY	2101	GCTCTGAACCTTCTCGGTTGAACCTGATTTAAAAACAAACATGATGACAGACCCGATTAC	2160
Db	2101	GCTCTGAACCTTCTCGGTTGAACCTGATTTAAAAACAAACATGATGACAGACCCGATTAC	2160
QY	2161	TCATTCGAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCATTCGAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
QY	2221	GCATGGTCCAAATCAGCACCCAGCAAAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC	2280
Db	2221	GCATGGTCCAAATCAGCACCCAGCAAAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC	2280
QY	2281	RAGGCTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAAACGGCCAGT	2340

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 14:24:04 ; Search time 1094.6 Seconds
(without alignments)
17846.836 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgacacatcgatgtgc.....acgacgttcaggactacgct 3300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	4 AAD11588	Aad11588 Partial s
2	3300	100.0	3300	4 AAD11589	Aad11589 Partial s
3	3226.4	97.8	3300	4 AAD11592	Aad11592 Partial s
4	3215.2	97.4	3300	4 AAD11591	Aad11591 Partial s
5	3207.2	97.2	3300	4 AAD11590	Aad11590 Partial s
6	2705.8	82.0	8101	4 AAD25681	Aad25681 Nucleotid
7	2705.8	82.0	8101	6 ABL53238	ABL53238 Nucleotid
8	2690.4	81.5	7259	12 ADO23608	Ado23608 DNA encod
9	2690.4	81.5	7322	12 ADO23613	Ado23613 DNA encod
10	2690.4	81.5	7352	12 ADO23603	Ado23603 DNA encod
11	2690.4	81.5	7370	12 ADO23598	Ado23598 DNA encod
12	2690.4	81.5	7370	12 ADO23639	Ado23639 DNA encod
13	2690.4	81.5	7370	12 ADO23649	Ado23649 DNA encod
14	2690.4	81.5	7370	12 ADO23644	Ado23644 DNA encod
15	2690.4	81.5	7403	12 ADO23594	Ado23594 DNA encod
16	2690.4	81.5	7442	12 ADO23590	Ado23590 DNA encod
17	2690.4	81.5	7478	12 ADO23584	Ado23584 DNA encod
18	2688.8	81.5	7112	12 ADO23618	Ado23618 DNA encod
19	2636.8	79.9	6806	6 ABL49925	ABL49925 Maltose b
20	2605.4	79.0	7553	12 ADO23588	Ado23588 DNA encod

21	1609.4	48.8	558	12	ADL90419	Adl90419 Clostridi
22	1474	44.7	4700	9	ACF06053	Acf06053 Vector pJ
23	1472.4	44.6	4700	12	ADL18670	Adl18670 Vector pJ
24	1462.4	44.3	4920	10	ADH73599	Adh73599 Plasmid p
25	1462.4	44.3	4935	10	ADH73598	Adh73598 Plasmid p
26	1462.4	44.3	4945	10	ADH73600	Adh73600 Plasmid p
27	1462.4	44.3	4951	10	ADH73601	Adh73601 Plasmid p
28	1418.2	43.0	5903	8	ABZ23939	Abz23939 Nucleotid
29	1417.6	43.0	1511	6	ABQ73200	Abq73200 E. coli t
C 30	1417.6	43.0	5926	2	AAV32977	Aav32977 Tn7 donor
C 31	1417.6	43.0	5926	6	AAD45059	Aad45059 Transposo
C 32	1417.6	43.0	5926	12	ADG46817	Adg46817 Donor pla
33	1362	41.3	5201	12	ADL72229	Adl72229 DNA seque
34	1362	41.3	5201	12	ADL72228	Adl72228 DNA seque
C 35	1238.4	37.5	5024	9	ACF06051	Acf06051 Plasmid p
C 36	1238.4	37.5	5024	12	ADL18668	Adl18668 Plasmid p
37	1237.2	37.5	1922	6	ABQ73089	Abq73089 MBP and z
38	1237.2	37.5	1922	10	ADF83622	Adf83622 MBP-human
39	1204	36.5	4557	2	AAT90491	Aat90491 Vector pl
40	1204	36.5	4969	12	ADO33522	Ado33522 Expressio
41	1204	36.5	5006	10	ADE24112	Ade24112 plasmid p
42	1204	36.5	5064	12	ADO33532	Ado33532 Vector pG
43	1204	36.5	5082	12	ADO33531	Ado33531 Vector pG
44	1204	36.5	6259	6	AAZ29720	Aaz29720 Plasmid p
45	1204	36.5	6823	3	AAC55459	Aac55459 Destinati

ALIGNMENTS

RESULT 1

AAD11588

ID AAD11588 standard; DNA; 3300 BP.

XX

AC AAD11588;

XX

DT 24-SEP-2001 (first entry)

XX

DE Partial sequence of pMal-m1 and ZFPm1 DNA.

XX

KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;

KW modulation; plant technology; agriculture; ds.

XX

OS Unidentified.

XX

PH Key

FT CDS

FT

FT Location/Qualifiers

2719..3270

/tag= a

/product= "ZFPm1 protein"

/transl_except= (pos:2785..2787, aa:Asp)

/transl_except= (pos:2788..2790, aa:Pro)

/transl_except= (pos:2791..2793, aa:Gly)

/transl_except= (pos:2872..2874, aa:Arg)

/transl_except= (pos:2875..2877, aa:Ala)

/transl_except= (pos:2878..2880, aa:His)

/transl_except= (pos:2884..2886, aa:Glu)

/transl_except= (pos:2953..2955, aa:Gln)

/transl_except= (pos:2959..2961, aa:Ser)

/transl_except= (pos:3121..3123, aa:Arg)

/transl_except= (pos:3124..3126, aa:Ser)

/transl_except= (pos:3127..3129, aa:Asp)

/transl_except= (pos:3130..3132, aa:Asn)

/transl_except= (pos:3136..3138, aa:Val)

/transl_except= (pos:3139..3141, aa:Arg)

/transl_except= (pos:3205..3207, aa:Gln)

/transl_except= (pos:3208..3210, aa:Ala)

/transl_except= (pos:3211..3213, aa:Gly)

/transl_except= (pos:3214..3216, aa:His)

/transl_except= (pos:3220..3222, aa:Ala)

/transl_except= (pos:3223..3225, aa:Ser)

/note= "CDS does not include start and stop codon"

/partial

2740..2790

primer_bind

```

primer_bind      /*tag= b  
/bound_moiety= "F1-f2 primer"  
2770.._2850  
/tag= c  
primer_bind      /bound_moiety= "F1-f1 primer"  
2824.._2889  
/tag= d  
primer_bind      /bound_moiety= "F2-b primer"  
2867.._2940  
/tag= e  
primer_bind      /bound_moiety= "F2-f primer"  
2916.._2973  
/tag= f  
primer_bind      /bound_moiety= "F3-b1 primer"  
2953.._3021  
/tag= g  
primer_bind      /bound_moiety= "F3-b2 primer"  
2992.._3042  
/tag= h  
primer_bind      /bound_moiety= "F4-f2 primer"  
3022.._3102  
/tag= i  
primer_bind      /bound_moiety= "F4-f1 primer"  
3076.._3141  
/tag= j  
primer_bind      /bound_moiety= "F5-b primer"  
3119.._3192  
/tag= k  
primer_bind      /bound_moiety= "F5-f primer"  
3168.._3225  
/tag= l  
primer_bind      /bound_moiety= "F6-b1 primer"  
3205.._3273  
/tag= m  
primer_bind      /bound_moiety= "F6-b2 primer"

```

WO200152620-A2.

26-JUL-2001.

19-JAN-2001; 2001WO-US001817.

21-JAN-2000: 2000US-0177468P.

44-38861-2000; 2000US-00620897.

(SYGN) SYNGENTA AGRIC DT

Barbara C.F. Steele, J.T. Green, Y. Dalmat

WPT: 2001-465335/50

P-PSDB; AAE06000.

New zinc finger p

plant cells a disorder that is associated with abnormal expression of the target gene.

Example 4; Page 138-139; 156pp: English.

The patent discloses methods and compositions to modulate the expression of a target gene in plant cells. The method involves providing plant cells with a zinc finger protein (ZFP) which is capable of specifically binding to a target nucleotide sequence or its complementary strand within a target gene and allowing the ZFP binding to the target nucleotide sequence, where the expression of the target gene in the plant cells is modulated. The ZFP and fusions of the ZFP proteins are useful for modulating or regulating gene expression and metabolic pathways in plants. The ZFP, fusion proteins and methods are useful in plant and agricultural technology. The method is useful particularly for treating a disorder in the plant cells, where the disorder is associated with abnormal expression of the target gene. The present DNA sequence is the partial sequence of pMal-m1 and ZFPm1 DNA

XX	SQ	Sequence	3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
		Query Match	100.0%; Score 3300; DB 4; Length 3300;
		Best Local Similarity	100.0%; Pred. No. 0;
		Matches 3300; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	1	CGGACACCATCGAATGGTGTGCAAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAAGAGA	60
Db	1	CGGACACCATCGAATGGTGTGCAAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAAGAGA	60
QY	61	GTCAATTCCAGGTGCTGAATGTGAAACCCAGTAAAGTTATACGATGTCGCAGAGTATGCCG	120
Db	61	GTCAATTCCAGGTGCTGAATGTGAAACCCAGTAAAGTTATACGATGTCGCAGAGTATGCCG	120
QY	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTGAACCCAGGCGAGCCACGTTTCTCGCAAAA	180
Db	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTGAACCCAGGCGAGCCACGTTTCTCGCAAAA	180
QY	181	CGCGGGAAAAGTCGGAACGGCGGATGGGAGCTGAATTACATTCCCAACCGCGTGGCAC	240
Db	181	CGCGGGAAAAGTCGGAACGGCGGATGGGAGCTGAATTACATTCCCAACCGCGTGGCAC	240
QY	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTGTGCAAAATTTGTTCGGCGGATTAATCTCGCGCGGATCAACTGGTGGCCAGCG	360
Db	301	ACGCGCGTGTGCAAAATTTGTTCGGCGGATTAATCTCGCGCGGATCAACTGGTGGCCAGCG	360
QY	361	TGTTGGTGTTCGATGTAGAACGAGCGGCTGCAGAGCTCTGAAGCGCGGTGCACAAATC	420
Db	361	TGTTGGTGTTCGATGTAGAACGAGCGGCTGCAGAGCTCTGAAGCGCGGTGCACAAATC	420
QY	421	TTTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACTATCGCGTGGATGACACGAGTCCCA	480
Db	421	TTTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACTATCGCGTGGATGACACGAGTCCCA	480
QY	481	TTGCTGTGGAGCTGCTGCACCTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAGA	540
Db	481	TTGCTGTGGAGCTGCTGCACCTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAGA	540
QY	541	CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTAGCGCACTGGGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTAGCGCACTGGGCGTGGAGCATC	600
QY	601	TGTTGCGCAATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660
Db	601	TGTTGCGCAATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660
QY	661	CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCGATAG	720
Db	661	CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCGATAG	720
QY	721	CGGAAACGGGAAGCGCATGAGTGGCATATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
Db	721	CGGAAACGGGAAGCGCATGAGTGGCATATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTTCGCCACTCGCATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA	840
Db	781	ATGAGGGCATCGTTCGCCACTCGCATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA	840
QY	841	TGGCGCGCATTTACGAGTCCGGCTCGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGGCGCGCATTTACGAGTCCGGCTCGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGTCTATGTTATATCCGCGGTTTAACCAACCATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGTCTATGTTATATCCGCGGTTTAACCAACCATCAACAGGATTTTC	960
QY	961	GCCTGTCTGGGGCAAAACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCGAGCGGTGA	1020

Db 961 GCCTGCTGGGCAAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAAGAAAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAAGAAAAACCAACCCCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTATGACAGTGGCAGCAGAGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTATGACAGTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGAAAGCGGCGAGTGAGGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG 1200
Qy 1201 GCACAATCTCATGTTTGAAGCTTATCATGCACTGCAAGCGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGAAGCTTATCATGCACTGCAAGCGTGCAACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGGCATCGGAAGCTGGTATGCTGTGACAGTCTGAGTCTGATCACTGATTAATTCG 1320
Db 1261 TCAGGAGGCATCGGAAGCTGGTATGCTGTGACAGTCTGAGTCTGATCACTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCATACGGTT 1380
Qy 1381 CTGGCAATPATTCTGAATGAGCTGTGTGAATAATATCATCGCTCGTATATGTGTGA 1440
Db 1381 CTGGCAATPATTCTGAATGAGCTGTGTGAATAATATCATCGCTCGTATATGTGTGA 1440
Qy 1441 ATTGTGAGGGATACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGGGATACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACATAGATTATGAAAACTGAAAGAGTTAAACTGGTAAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACATAGATTATGAAAACTGAAAGAGTTAAACTGGTAAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAGTTCGGAAGTTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAAAGCTCTCGCTGAAGTTCGGAAGTTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGCTGAGCATCCGATTAACCTGGAAGAGAAATCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGCTGAGCATCCGATTAACCTGGAAGAGAAATCCACAGGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAACAGCCGCTTCTGGTGTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAACAGCCGCTTCTGGTGTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT 1800
Qy 1801 CGGTTTACCTGGATCGCGTACGTTTCAACCGGCAAGCTGATTGCTTTACCGGATCGCTGTT 1860
Db 1801 CGGTTTACCTGGATCGCGTACGTTTCAACCGGCAAGCTGATTGCTTTACCGGATCGCTGTT 1860
Qy 1861 GAAGCGTATCGTGAATTTATACAAAGATCTGCTGCCAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTATCGTGAATTTATACAAAGATCTGCTGCCAACCCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGATTAAGAACTGAAAGCGAAAGGTAAGAGCGCGTGTATTCAC 1980
Db 1921 GAGATCCCGCGCTGATTAAGAACTGAAAGCGAAAGGTAAGAGCGCGTGTATTCAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGCGCTGATTGCTGTGAACGGGGTTATGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCGCTGATTGCTGTGAACGGGGTTATGGTTCAAG 2040
Qy 2041 TATGAAACCGGCAAGTACGACATTAAGACGTTGGGCTGGATTAACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTAAGACGTTGGGCTGGATTAACGCTGGCGGAAGCG 2100

Qy 2101 GGTCTGACCTTTCCTGGTTGACCTGATTAATAAACAACAATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTTCCTGGTTGACCTGATTAATAAACAACAATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGCAACAGCATGACCATCAAGCGGCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGCAACAGCATGACCATCAAGCGGCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGGTAAACGGTACTGCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGGTAAACGGTACTGCGACCTTC 2280
Qy 2281 AAGGTTCAACCATCAACACCGTTTGGCGTGTGAGCGCAGGTATTAAGCGGCCAGT 2340
Db 2281 AAGGTTCAACCATCAACACCGTTTGGCGTGTGAGCGCAGGTATTAAGCGGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGGAAAGATCCACGTTATTTGGCGCCACCATGMAAACGCCACAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTTATTTGGCGCCACCATGMAAACGCCACAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGCTTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTCTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACAGTCTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGAGAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGAGAGGATTCAGAAATTC 2700
Qy 2701 GGATCTCTCTCTGTCGGCCAGCGCCCTCGAGCCCGGGAGAACCCCTATCTCTTGT 2760
Db 2701 GGATCTCTCTCTGTCGGCCAGCGCCCTCGAGCCCGGGAGAACCCCTATCTCTTGT 2760
Qy 2761 CCGGAATGTGGTAACTCTCTCTCAGAGCTCTCACCTGGTGGCCACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAACTCTCTCTCAGAGCTCTCACCTGGTGGCCACAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGACTCATATCGCGAGAGCCATACAAATGTCCAGAAATGTGCG 2940
Db 2881 CTGGTGGCCATCAACGACTCATATCGCGAGAGCCATACAAATGTCCAGAAATGTGCG 2940
Qy 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGCGACCAACGTTACTCACACGGGAGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGCGACCAACGTTACTCACACGGGAGAAG 3000
Qy 3001 CCTATGCTTGTCCGGAATGTGGTAACTTTAGTCTTTAGCCGAGCGATTAACCTGGTGGCCAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGGTAACTTTAGTCTTTAGCCGAGCGATTAACCTGGTGGCCAC 3060
Qy 3061 CAGCGTACCCACCGGGTGAATAACCGGTATAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACCGGGTGAATAACCGGTATAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Qy 3121 CAGGCGGCGCACTGGCCAGCCATCAACGCACTCATCTGGCGAGAGCCATACAAATGT 3180
Db 3121 CAGGCGGCGCACTGGCCAGCCATCAACGCACTCATCTGGCGAGAGCCATACAAATGT 3180

QY 3181 CCAGAAATGGCAAGTCTTCTCTCGGCTGCAAAATCTCGTCCGGCACCACAGTACTCAC 3240
 Db |||||
 3181 CCAGAAATGGCAAGTCTTCTCTCGGCTGCAAAATCTCGTCCGGCACCACAGTACTCAC 3240
 QY 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCGAGTACCGGTACGAGTTCCGGACTACGCT 3300
 Db |||||
 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCGAGTACCGGTACGAGTTCCGGACTACGCT 3300

RESULT 2

AAD11589

ID AAD11589 standard; DNA; 3300 BP.

XX AAD11589;

XX 24-SEP-2001 (first entry)

XX Partial sequence of pMal-m2 and ZFPm2 DNA.

XX Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
 KW modulation; plant technology; agriculture; ds.
 XX Unidentified.

OS Unidentified.

PH Key Location/Qualifiers

FT CDS 2719..3270

FT /tag= a

FT /product= "ZFPm2 protein"

FT /note= "CDS does not include start and stop codon"

FT /partial

FT primer_bind 2740..2790

FT /tag= b

FT /bound_moiety= "F1-f2 primer"

FT primer_bind 2770..2850

FT /tag= c

FT /bound_moiety= "F1-f1 primer"

FT primer_bind 2824..2889

FT /tag= d

FT /bound_moiety= "F2-b primer"

FT primer_bind 2867..2940

FT /tag= e

FT /bound_moiety= "F2-f primer"

FT primer_bind 2916..2973

FT /tag= f

FT /bound_moiety= "F3-b1 primer"

FT primer_bind 2953..3021

FT /tag= g

FT /bound_moiety= "F3-b2 primer"

FT primer_bind 2992..3042

FT /tag= h

FT /bound_moiety= "F4-f2 primer"

FT primer_bind 3022..3102

FT /tag= i

FT /bound_moiety= "F4-f1 primer"

FT primer_bind 3076..3141

FT /tag= j

FT /bound_moiety= "F5-b primer"

FT primer_bind 3119..3192

FT /tag= k

FT /bound_moiety= "F5-f primer"

FT primer_bind 3168..3225

FT /tag= l

FT /bound_moiety= "F6-b1 primer"

FT primer_bind 3205..3273

FT /tag= m

FT /bound_moiety= "F6-b2 primer"

XX WO200152620-A2.

XX 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US001817.

XX

PR 21-JAN-2000; 2000US-0177468P.
 PR 21-JUL-2000; 2000US-00620897.
 XX (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA AGRIC DISCOVERY INC.
 XX Barbas CF, Stege JT, Guan X, Dalmia B;
 XX WPI; 2001-465325/50.
 DR P-FSDB; AAE06002.
 XX
 PT New zinc finger proteins, useful for modulating or regulating gene
 PT expression and metabolic pathways in plants, e.g. for treating in the
 PT plant cells a disorder that is associated with abnormal expression of the
 PT target gene.
 XX
 PS Example 4; Page 140-142; 156pp; English.
 CC The patent discloses methods and compositions to modulate the expression
 CC of a target gene in plant cells. The method involves providing plant
 CC cells with a zinc finger protein (ZFP) which is capable of specifically
 CC binding to a target nucleotide sequence or its complementary strand
 CC within a target gene and allowing the ZFP binding to the target
 CC nucleotide sequence, where the expression of the target gene in the plant
 CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
 CC for modulating or regulating gene expression and metabolic pathways in
 CC plants. The ZFP, fusion proteins and methods are useful in plant and
 CC agricultural technology. The method is useful particularly for treating a
 CC disorder in the plant cells, where the disorder is associated with
 CC abnormal expression of the target gene. The present DNA sequence is the
 CC partial sequence of pMal-m2 and ZFPm2 DNA
 XX
 SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
 Query Match 100.0%; Score 3300; DB 4; Length 3300;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAGAGA 60
 Db |||||
 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAGAGA 60
 QY 61 GTCAATTACAGGTTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
 Db |||||
 61 GTCAATTACAGGTTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
 QY 121 GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACAGCGCCAGCGTTCCTGCGAAA 180
 Db |||||
 121 GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACAGCGCCAGCGTTCCTGCGAAA 180
 QY 181 CGCGGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGGTGGCAC 240
 Db |||||
 181 CGCGGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGGTGGCAC 240
 QY 241 AACAACTGGCGGCAAAACAGTCTGTTGATGGCGTTGCCACTCCAGTCTGGCCCTGC 300
 Db |||||
 241 AACAACTGGCGGCAAAACAGTCTGTTGATGGCGTTGCCACTCCAGTCTGGCCCTGC 300
 QY 301 ACGCGCGTGCAGAAATTTGTCGCGGATTAATACTCGCGCGATCACTGGGTGCCAGCG 360
 Db |||||
 301 ACGCGCGTGCAGAAATTTGTCGCGGATTAATACTCGCGCGATCACTGGGTGCCAGCG 360
 QY 361 TGGTGGTGTGATGGTAGAACGAAAGCGCGTCAAGCCTGTAAGCGCGGTGCACAATC 420
 Db |||||
 361 TGGTGGTGTGATGGTAGAACGAAAGCGCGTCAAGCCTGTAAGCGCGGTGCACAATC 420
 QY 421 TTCTCGCGCAACCGGTGAGTGGCTGATCATTAATTAATCGGTGGATGACCAAGATGCCA 480
 Db |||||
 421 TTCTCGCGCAACCGGTGAGTGGCTGATCATTAATTAATCGGTGGATGACCAAGATGCCA 480
 QY 481 TTGCTGTGGAAGCTGCTGCACATAATTTCCGCGGTTATTTCTTGATGTCCTGACGAGA 540
 Db |||||
 481 TTGCTGTGGAAGCTGCTGCACATAATTTCCGCGGTTATTTCTTGATGTCCTGACGAGA 540

QY	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGTGGAGCATC	600
QY	601	TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCCCCATTAAGTTCTGTCGCG	660
DB	601	TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCCCCATTAAGTTCTGTCGCG	660
QY	661	CGCGTCTGCGTCTGGCTGGCTGGCTGAATAATCTCACTCGCAATCAAAATTCAGCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGCTGGCTGAATAATCTCACTCGCAATCAAAATTCAGCGATAG	720
QY	721	CGGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCTTCCCACTGGATGCTGGTTGCCAAGCATCAGATGGCGTGGCGGCA	840
DB	781	ATGAGGGCATCTTCCCACTGGATGCTGGTTGCCAAGCATCAGATGGCGTGGCGGCA	840
QY	841	TGCGGCCATTACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGGCCATTACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAAACCAATCAAAAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAAACCAATCAAAAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020
DB	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTGGCCGCTCACTGCTGAAAGAAACCAACCTGGCGGCCAATA	1080
DB	1021	AGGGCAATCAGCTGTGGCCGCTCACTGCTGAAAGAAACCAACCTGGCGGCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCACTGGCAGCACAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCACTGGCAGCACAGGTTT	1140
QY	1141	CCGACTGAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCATTTAG	1200
DB	1141	CCGACTGAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGACGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGCGAGCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATATTCG	1320
DB	1261	TCAGGCGAGCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATATTCG	1320
QY	1321	TGTGCTCAAGCGGCACTCCGTTCTGGATATGTTTTTTTGGCCGACATCAACGGTT	1380
DB	1321	TGTGCTCAAGCGGCACTCCGTTCTGGATATGTTTTTTTGGCCGACATCAACGGTT	1380
QY	1381	CTGGCAATATTTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATATGTTGGA	1440
DB	1381	CTGGCAATATTTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTCACAGGAACAGGCGAGTCCGTTTGGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATAACAAATTCACAGGAACAGGCGAGTCCGTTTGGTGTTCACGA	1500
QY	1501	GCACTTCACCAACAGGACCATAGATTATGAATACTGAAGAGGTAAACTGGTAATCTGG	1560
DB	1501	GCACTTCACCAACAGGACCATAGATTATGAATACTGAAGAGGTAAACTGGTAATCTGG	1560
QY	1561	ATTACGGGCAATTAAGGCTATTAACGGTCTGCTGTAAGTTCGGTAAGAAATTCGAAAGAT	1620
DB	1561	ATTACGGGCAATTAAGGCTATTAACGGTCTGCTGTAAGTTCGGTAAGAAATTCGAAAGAT	1620

QY	1621	ACCGAATTAAGTCAACGGTTGAGCATCCGATAAACTGGGAAGAAATTTCCACAGGTT	1680
DB	1621	ACCGAATTAAGTCAACGGTTGAGCATCCGATAAACTGGGAAGAAATTTCCACAGGTT	1680
QY	1681	CGCGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
DB	1681	CGCGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCGCTGTGGCTGGAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGAT	1800
DB	1741	GCTCAATCTGCGCTGTGGCTGGAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGAT	1800
QY	1801	CGGTTACTCGGATGGCTAGCTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT	1860
DB	1801	CGGTTACTCGGATGGCTAGCTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT	1860
QY	1861	GAAGGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
DB	1861	GAAGGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTAAGAGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTAAGAGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAACGCGCACTACGACATTAACAGCTGGCGCTGGATAACGCTGGCGCGAAAGCG	2100
DB	2041	TATGAAAACGCGCACTACGACATTAACAGCTGGCGCTGGATAACGCTGGCGCGAAAGCG	2100
QY	2101	GCTCTGACCTTCTCGTTGACCTGATTAACAAACACATGATGACAGACACCGATTAC	2160
DB	2101	GCTCTGACCTTCTCGTTGACCTGATTAACAAACACATGATGACAGACACCGATTAC	2160
QY	2161	TCCATCGAGAAGCTGCGCTTTAATAAGCGAAGACAGCGATGACCATCAACGGCCGCTGG	2220
DB	2161	TCCATCGAGAAGCTGCGCTTTAATAAGCGAAGACAGCGATGACCATCAACGGCCGCTGG	2220
QY	2221	GCATGGTCCACATCGACACAGCAAGTGAATTAATGTTAAACGCTACGCGACCTTC	2280
DB	2221	GCATGGTCCACATCGACACAGCAAGTGAATTAATGTTAAACGCTACGCGACCTTC	2280
QY	2281	AAGGCTCAACCATCAAAACCGTTCTGCGCTGCTGAGCGCAGGTATTAACCGCCAGT	2340
DB	2281	AAGGCTCAACCATCAAAACCGTTCTGCGCTGCTGAGCGCAGGTATTAACCGCCAGT	2340
QY	2341	CGAAACAAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG	2400
DB	2341	CGAAACAAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG	2400
QY	2401	GNAAGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACCAGGAAGAG	2460
DB	2401	GNAAGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACCAGGAAGAG	2460
QY	2461	TTGGCGAAAGATCCACGTATTTCGCCCACTGGAACCGCCAGAAAGGTGAATTCATG	2520
DB	2461	TTGGCGAAAGATCCACGTATTTCGCCCACTGGAACCGCCAGAAAGGTGAATTCATG	2520
QY	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGTGCCTGCTGCTGATCAACGCG	2580
DB	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGTGCCTGCTGCTGATCAACGCG	2580
QY	2581	GCCAGGGTCTGACATGTCGATGAAGCCCTGAAAGACCGCCAGACTTAATTCGAGCTCG	2640
DB	2581	GCCAGGGTCTGACATGTCGATGAAGCCCTGAAAGACCGCCAGACTTAATTCGAGCTCG	2640
QY	2641	AACAACAAACAATAACAAATTAACAAACCTCGGGATCGAGGAGGAGTTTCAGAAATTC	2700
DB	2641	AACAACAAACAATAACAAATTAACAAACCTCGGGATCGAGGAGGAGTTTCAGAAATTC	2700
QY	2701	GGATCCTCTTCTGCTGGCCCGAGCGGCTTCGAGCCCGGGGAGAAAGCCCTATGCTTGT	2760


```
Db 2701 GGATCCTCTCTCTGTGGCCAGCGGCGCTCGAGCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTACCTGGTGGCCACGAGGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTACCTGGTGGCCACGAGGTACCCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTGGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTGGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCATCAACGCACTCATCTCGTCCGGCACCACGTAACATGTCGAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATCTCGTCCGGCACCACGTAACATGTCGAGATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCACGTAACATGTCGAGATGTGGC 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCACGTAACATGTCGAGATGTGGC 3000
QY 3001 CCCTATGTTGTCGGAAATGTGGTAAGTCTTTCAGCGCAGCGATAACCTGGTGGCCAC 3060
Db 3001 CCCTATGTTGTCGGAAATGTGGTAAGTCTTTCAGCGCAGCGATAACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTGGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTGGC 3120
QY 3121 CAGCGCGCCACCTGGCCAGCCATCAACGCACTCATCTGGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTGGCCAGCCATCAACGCACTCATCTGGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAATGTGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCACGTAACCTAC 3240
Db 3181 CCAGAATGTGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCACGTAACCTAC 3240
QY 3241 ACCGTTAAAAAAGTGGCCAGCGCCGCGCAGTACCGTACCGTACCGTACCGTACCGT 3300
Db 3241 ACCGTTAAAAAAGTGGCCAGCGCCGCGCAGTACCGTACCGTACCGTACCGTACCGT 3300

RESULT 3
ID AAD11592 standard; DNA; 3300 BP.
XX
AC AAD11592;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-Ap3 and ZFPap3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; Ap3; APETALA3; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2719..3270
FT /tag= a
FT /product= "ZFPap3 protein"
FT /transl_except= (pos:3124..3126, aa:Ser)
FT /transl_except= (pos:3127..3129, aa:Ser)
FT /transl_except= (pos:3136..3138, aa:Val)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT
```

```
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3169..3225
FT /tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /tag= m
FT /bound_moiety= "F6-b2 primer"
XX WO200152620-A2.
PN 26-JUL-2001.
XX 19-JAN-2001; 2001WO-US001817.
XX 21-JAN-2000; 2000US-0177488P.
XX 21-JUL-2000; 2000US-00620897.
XX (SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI Barbas CP, Stege JT, Guan X, Dalmia B;
XX WPI; 2001-465325/50.
DR P-PSDB; AAE06005.
XX
PT New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 148-149; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-Ap3 (APETALA3) and ZFPap3 DNA
XX
SQ Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;
Query Match 97.8%; Score 3226.4; DB 4; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGTGTGCAAAACCTTTCGGGTATGGCATGTAGGCCCGGAGAGA 60
```

Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGGCGGTATGGCATGATAGGCCCGGAAGAGA 60
Qy 61 GTCAATTGAGGTGTGAATGTGAACACAGTAACCGTTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTGAGGTGTGAATGTGAACACAGTAACCGTTATACGATGTGCGAGAGTATGCCG 120
Qy 121 GTGTCCTCTATCAGACCGTTTCCCGCGTGGTGAACCGGCCAGCCAGCTTTCTGCGAANA 180
Db 121 GTGTCCTCTATCAGACCGTTTCCCGCGTGGTGAACCGGCCAGCCAGCTTTCTGCGAANA 180
Qy 181 CCGGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Db 181 CCGGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Qy 241 AACAACTCGCGGGCAAAACAGTGTGCTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
Db 241 AACAACTCGCGGGCAAAACAGTGTGCTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
Qy 301 AGCGCCGTGCGCAAAATGTGCGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 360
Db 301 AGCGCCGTGCGCAAAATGTGCGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 360
Qy 361 TGGTGTGTGCAATGTAGAACGAGCGCGTGAAGCGCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGTGTGCAATGTAGAACGAGCGCGTGAAGCGCTGTAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACGCGTCAAGTGGGTGATCATTAACATATCCGCTGGATGACCAAGTGC 480
Db 421 TTCTCGCGCAACGCGTCAAGTGGGTGATCATTAACATATCCGCTGGATGACCAAGTGC 480
Qy 481 TTGCTGTGGAAGCTGCCTGCAATATGTTCCGGCGTATTTCTTGATGCTCTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCAATATGTTCCGGCGTATTTCTTGATGCTCTGACACAGA 540
Qy 541 CACCCATCAACAGATATTATTTCTCCCATGAAGCGGTACGGACTGGCGGTGAAGCATC 600
Db 541 CACCCATCAACAGATATTATTTCTCCCATGAAGCGGTACGGACTGGCGGTGAAGCATC 600
Qy 601 TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCG 660
Db 601 TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Qy 721 CCGAACGGGAAGCGACCTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAACGGGAAGCGACCTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Qy 841 TCGCGGCCATTAACGAGTCCGGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCCATTAACGAGTCCGGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACACGAGCTGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAACACGAGCTGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCAGTGGCAGCAGAGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCAGTGGCAGCAGAGTTT 1140

Db 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCAGTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCCAAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCCAAATTAATGTAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATTCATGTTTGACAGCTTATCATCGATGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTCATGTTTGACAGCTTATCATCGATGCAAGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTGTGAGGTGTCAGGTGCTTAATCATCTGCATTAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTGTGAGGTGTCAGGTGCTTAATCATCTGCATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTCGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTCGCGGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTCG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTCG 1560
Qy 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATATACTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATATACTGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 CGCGCAACTGCGGATGCGCTGACATTAATCTTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGCGGATGCGCTGACATTAATCTTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGCAAAAGGTTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGCAAAAGGTTTCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAAGGCAAGCTGATTGCTTACCGCATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAAGGCAAGCTGATTGCTTACCGCATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCAAGATAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCAAGATAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCCACCTGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATAACGCTGGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATAACGCTGGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATGAAATGCAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATGAAATGCAGACACCGGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGGAAACAGCGATGACCATCAAGGCGCCGTTG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGGAAACAGCGATGACCATCAAGGCGCCGTTG 2220

```
QY 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTATGTGTAAACGGTACTGCGGACCTTC 2280
Db |||
QY 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTATGTGTAAACGGTACTGCGGACCTTC 2280
Db |||
QY 2281 AAGGTCACACCTCAAAACCGTTCGTTGGCTGCTGAGCGGAGGATTAACGCGGCCAGT 2340
Db |||
QY 2281 AAGGTCACACCTCAAAACCGTTCGTTGGCTGCTGAGCGGAGGATTAACGCGGCCAGT 2340
Db |||
QY 2341 CGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db |||
QY 2341 CGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db |||
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db |||
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db |||
QY 2461 TTGGCGAAGATCCAGCTATTGCGGCCACCATGGAAGACGCCGCAAGAGTGAATCATG 2520
Db |||
QY 2461 TTGGCGAAGATCCAGCTATTGCGGCCACCATGGAAGACGCCGCAAGAGTGAATCATG 2520
Db |||
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCGGTGCTGCTGCGGTGATCAAGCC 2580
Db |||
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCGGTGCTGCTGCGGTGATCAAGCC 2580
Db |||
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db |||
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db |||
QY 2641 AACACACACACATTAACAATAACAACCTCCGGATCGAGGAGGATTCAGAAATTC 2700
Db |||
QY 2641 AACACACACACATTAACAATAACAACCTCCGGATCGAGGAGGATTCAGAAATTC 2700
Db |||
QY 2701 GGATCCTCTTCTGTCGCGCCAGCGCCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db |||
QY 2701 GGATCCTCTTCTGTCGCGCCAGCGCCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db |||
QY 2761 CCGGAATGTGTAAGTCTTCTCAGAGCTCTCAGAGTGTGCGCCACACGAGCTACCCAC 2820
Db |||
QY 2761 CCGGAATGTGTAAGTCTTCTCAGAGCTCTCAGAGTGTGCGCCACACGAGCTACCCAC 2820
Db |||
QY 2821 ACGGTGAAACACCGTAAATATGCCAGAGTGGCGCAATCTTTAGCAGTCCAGCAAC 2880
Db |||
QY 2821 ACGGTGAAACACCGTAAATATGCCAGAGTGGCGCAATCTTTAGCAGTCCAGCAAC 2880
Db |||
QY 2881 CTGGTGCGCCATCAACGACCTCATACTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db |||
QY 2881 CTGGTGCGCCATCAACGACCTCATACTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db |||
QY 2941 AAGTCTTTCTCGGTCTGACAAATCTCGTCCGGCACCAAGTACTCACCGGGGAGAAG 3000
Db |||
QY 2941 AAGTCTTTCTCGGTCTGACAAATCTCGTCCGGCACCAAGTACTCACCGGGGAGAAG 3000
Db |||
QY 3001 CCCTATGCTTCTCGGAGTGGTAAGTCTTCTAGCGCGAGGATACCTGGTGGCGCAC 3060
Db |||
QY 3001 CCCTATGCTTCTCGGAGTGGTAAGTCTTCTAGCGCGAGGATACCTGGTGGCGCAC 3060
Db |||
QY 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCGCGAGTGGCGCAATCTTTAGC 3120
Db |||
QY 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCGCGAGTGGCGCAATCTTTAGC 3120
Db |||
QY 3121 CAGCGCGCCACCTGGCGAGCCATCAACGCACTCATCTGGCGAGAGCCATACAAATGT 3180
Db |||
QY 3121 CAGCGCGCCACCTGGCGAGCCATCAACGCACTCATCTGGCGAGAGCCATACAAATGT 3180
Db |||
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCTGCGGACCAACGCTACTCAC 3240
Db |||
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCTGCGGACCAACGCTACTCAC 3240
Db |||
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTAACGAGTTCGGGACTAGCT 3300
Db |||
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTAACGAGTTCGGGACTAGCT 3300
Db |||
```

```
RESULT 4
AAD11591
ID AAD11591 standard; DNA; 3300 BP.
XX
AC AAD11591;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m4 and ZFPm4 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2719..3270
FT /tag= a
FT /product= "ZFPm4 protein"
FT /transl_except= (pos:3046..3048, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /tag= m
FT /bound_moiety= "F6-b2 primer"
XX
PN WO200152620-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US001817.
XX
PR 21-JAN-2000; 2000US-0177468P.
XX
PR 21-JUL-2000; 2000US-0062089P.
XX
PA (SCRI ) SCRIPPS RES INST.
XX
PA (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI Barbas CF, Stege JT, Guan X, Dalmia B;
```



```
QY 1741 GCTCAATCTGGCTGTTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGCTGATCAACCGGAAAGCTGATGCTTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGCTGATCAACCGGAAAGCTGATGCTTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCCTGCTGATTTAACAAGATCTGCTGCGGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCCTGCTGATTTAACAAGATCTGCTGCGGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCGGAGTAAAGCGCGCTGATGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCGGAGTAAAGCGCGCTGATGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGGTGATTAAGCGCGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGGTGATTAAGCGCGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGATTAAGAAACCAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGATTAAGAAACCAACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGGAGAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAACGGGCGCGTGG 2220
Db 2161 TCCATCGGAGAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAACGGGCGCGTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAAGTAAACGGTACTCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAAGTAAACGGTACTCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAACACCGTTCTGCTGGCGTGAAGTAAACGGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAACACCGTTCTGCTGGCGTGAAGTAAACGGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGATGAAGGCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGATGAAGGCTG 2400
QY 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAAG 2460
Db 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATGCGGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATGCGGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGCTCGCTTTCTGCTGATGCGGTGCGTACTCGGCTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGCTCGCTTTCTGCTGATGCGGTGCGTACTCGGCTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGGCAGACTAAATTCGAGTCTG 2640
QY 2641 AACAAACAACAATTAACAATTAACAACCTCGGGATCGAGGAAAGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATTAACAATTAACAACCTCGGGATCGAGGAAAGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGCTGGCCAGCGGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGCTGGCCAGCGGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAAATGTGTAAGTCTTCTCTCAGAGCTCTCAGCTGCTGCGCCACAGCGTACCCAC 2820
Db 2761 CCGGAAATGTGTAAGTCTTCTCTCAGAGCTCTCAGCTGCTGCGCCACAGCGTACCCAC 2820
```

```
QY 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCAGTCCAGCAAC 2880
Db 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCAGTCCAGCAGC 2880
QY 2881 CTGGTGGCCCATCAACGCACTCATCTGGCGAGAAAGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATCTGGCGAGAAAGCCATACAAATGTCCAGAAATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTCACACGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTCACACGGGGAGAAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCAGAGCTCTCACCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCAGAGCTCTCACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
QY 3121 CAGCCCGGCACCTGGCCAGCCATCAACGCACTCATCTGCGGAGAAAGCCATACAAATGT 3180
Db 3121 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCCAGTACCCGTACGACGTTCCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCCAGTACCCGTACGACGTTCCGGGACTACGCT 3300

RESULT 5
AAD11590
ID AAD11590 standard; DNA; 3300 BP.
XX
AC AAD11590;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m3 and ZFPm3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2719..3270
FT /tag= a
FT /product= "ZFPm3 protein"
FT /note= "CDS does not include start and stop codon"
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
```

```
FT      /*tag= h
FT      /bound_moiety= "F4-f2 primer"
FT      3022..3102
FT      /*tag= i
FT      /bound_moiety= "F4-f1 primer"
FT      3076..3141
FT      /*tag= j
FT      /bound_moiety= "F5-b primer"
FT      3119..3192
FT      /*tag= k
FT      /bound_moiety= "F5-f primer"
FT      3168..3225
FT      /*tag= l
FT      /bound_moiety= "F6-b1 primer"
FT      3205..3273
FT      /*tag= m
FT      /bound_moiety= "F6-b2 primer"
XX
PN      WO200152620-A2.
XX
XX      26-JUL-2001.
PD
XX      19-JAN-2001; 2001WO-US001817.
XX
XX      21-JAN-2000; 2000US-0177468P.
PR
XX      21-JUL-2000; 2000US-00620897.
PR
XX
PA      (SCRI ) SCRIPPS RES INST.
PA      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX      Barbas CF, Stege JT, Guan X, Dalmia B;
PI
XX      WPI; 2001-465325/50.
DR
XX      P-PSDB; AA06003.
XX
XX      New zinc finger proteins, useful for modulating or regulating gene
PT      expression and metabolic pathways in plants, e.g. for treating in the
PT      plant cells a disorder that is associated with abnormal expression of the
PT      target gene.
XX
PS      Example 4; Page 143-145; 156pp; English.
XX
XX      The patent discloses methods and compositions to modulate the expression
CC      of a target gene in plant cells. The method involves providing plant
CC      cells with a zinc finger protein (ZFP) which is capable of specifically
CC      binding to a target nucleotide sequence or its complementary strand
CC      within a target gene and allowing the ZFP binding to the target
CC      nucleotide sequence, where the expression of the target gene in the plant
CC      cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC      for modulating or regulating gene expression and metabolic pathways in
CC      plants. The ZFP, fusion proteins and methods are useful in plant and
CC      agricultural technology. The method is useful particularly for treating a
CC      disorder in the plant cells, where the disorder is associated with
CC      abnormal expression of the target gene. The present DNA sequence is the
CC      Partial sequence of pMal-m3 and ZFPm3 DNA
XX
SQ      Sequence 3300 BP; 845 A; 878 C; 877 G; 700 T; 0 U; 0 Other;
      Query Match      97.2%; Score 3207.2; DB 4; Length 3300;
      Best Local Similarity 98.2%;      P-PSDB; AA06003.
      Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
      1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA 60
      1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA 60
      61 GTCAATTCAGGTGGTGAATGTGAACCAAGTAAACGTTATACGATTCGCAGAGTATCGCG 120
      61 GTCAATTCAGGTGGTGAATGTGAACCAAGTAAACGTTATACGATTCGCAGAGTATCGCG 120
      121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCCAGCCAGCTTCTCGGAAA 180
      121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCCAGCCAGCTTCTCGGAAA 180
```

```
QY      181 CGCGGGAAGAGTGGAAAGCGCGCATGGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCAC 240
DB      |||||
DB      181 CGCGGGAAGAGTGGAAAGCGCGCATGGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCAC 240
QY      241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
DB      |||||
DB      241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY      301 ACGCCGCGTCGCAAAATTTGTCGCGCGGATTAATATCTCGCGCGGATCAACTGGGTGCGACG 360
DB      |||||
DB      301 ACGCCGCGTCGCAAAATTTGTCGCGCGGATTAATATCTCGCGCGGATCAACTGGGTGCGACG 360
QY      361 TGGTGGTGTGATGCTAGAACGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGACAAATC 420
DB      |||||
DB      361 TGGTGGTGTGATGCTAGAACGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGACAAATC 420
QY      421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCGCTGATCGCTGATGACCAAGATGCCA 480
DB      |||||
DB      421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCGCTGATCGCTGATGACCAAGATGCCA 480
QY      481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTTGATGCTCTGACCCAGA 540
DB      |||||
DB      481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTTGATGCTCTGACCCAGA 540
QY      541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACCGGACTGGGGCTGGAGCATC 600
DB      |||||
DB      541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACCGGACTGGGGCTGGAGCATC 600
QY      601 TGGTCGCAATTTGGGTCCACGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGCTCGG 660
DB      |||||
DB      601 TGGTCGCAATTTGGGTCCACGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGCTCGG 660
QY      661 CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB      |||||
DB      661 CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY      721 CGGAACGGGAAGGCGACATGGAGTGCATGTCGGTTCGGTTTCAACAAACCATGCAAAATGCTGA 780
DB      |||||
DB      721 CGGAACGGGAAGGCGACATGGAGTGCATGTCGGTTCGGTTCGGTTTCAACAAACCATGCAAAATGCTGA 780
QY      781 ATGAGGCGATCGTTCCCACTCGCATGCTGGTTCGCAACGATGATGCGGTGGCGGCA 840
DB      |||||
DB      781 ATGAGGCGATCGTTCCCACTCGCATGCTGGTTCGCAACGATGATGCGGTGGCGGCA 840
QY      841 TGGCGCCATTACCGAGTCCCGGCTGCGGCTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB      |||||
DB      841 TGGCGCCATTACCGAGTCCCGGCTGCGGCTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY      901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCAATCAAAACAGGATTTTC 960
DB      |||||
DB      901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCAATCAAAACAGGATTTTC 960
QY      961 GCCTGCTGGGCAAAACGATGCGCTGCGTTCGTCGAATCTCTCAGGCGCCAGGCGGTGA 1020
DB      |||||
DB      961 GCCTGCTGGGCAAAACGATGCGCTGCGTTCGTCGAATCTCTCAGGCGCCAGGCGGTGA 1020
QY      1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGCTGAAAGAAAACCAACCTGGCGCCCAATA 1080
DB      |||||
DB      1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGCTGAAAGAAAACCAACCTGGCGCCCAATA 1080
QY      1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCATTAAATGACAGTGGCAGACAGGTTT 1140
DB      |||||
DB      1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCATTAAATGACAGTGGCAGACAGGTTT 1140
QY      1141 CCGCATGGAAGCGGCGAGTGGCGCAACCGCAATTAATGAGTACTCACTCATTTAG 1200
DB      |||||
DB      1141 CCGCATGGAAGCGGCGAGTGGCGCAACCGCAATTAATGAGTACTCACTCATTTAG 1200
QY      1201 GCACAATTTCTATGTTTGACAGCTTATCATGACCTGACGCTGCAACCAATGCTTCTGGCG 1260
DB      |||||
DB      1201 GCACAATTTCTATGTTTGACAGCTTATCATGACCTGACGCTGCAACCAATGCTTCTGGCG 1260
```

```
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATATTCG 1320
QY 1321 TGTGCGTCAAGGGGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTGCGTCAAGGGGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTACACAGGAACAGCCAGTCGGTTTAGTGTTTTACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAACAGCCAGTCGGTTTAGTGTTTTACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTCGGTAACTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTCGGTAACTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT 1680
QY 1681 GCGCAATGCGGATGGCCCTGACATTAATCTTGGGCACAGCAGCGTTTGGTGCTAC 1740
Db 1681 GCGCAATGCGGATGGCCCTGACATTAATCTTGGGCACAGCAGCGTTTGGTGCTAC 1740
QY 1741 GCTCAATCTGGGCTGTGGCTGAATATCAACCGGACAAAGCGTTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGGCTGTGGCTGAATATCAACCGGACAAAGCGTTTCCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTAAACCGCAAGCTGATGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTAAACCGCAAGCTGATGCTTACCGGATCGCTGTT 1860
QY 1861 GAAGGTTATCGCTGATTTATAAAGATCTGTGCGGACCGGACCGGCAAAACCTGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAAAGATCTGTGCGGACCGGACCGGCAAAACCTGGAA 1920
QY 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCTGACTTCACTGCGCGCTGATGCTGCTGACGGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCTGACTTCACTGCGCGCTGATGCTGCTGACGGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATACGCTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACACATGAATGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGTTGTAAGCTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGTTGTAAGCTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCAATCAACCGTTCTGGTGGCTGTGAGCGCAGGATTAACCGCGCCAGT 2340
Db 2281 AAGGGTCAACCAATCAACCGTTCTGGTGGCTGTGAGCGCAGGATTAACCGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
```

```
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACCTATCTGCTGACTGATGAAGGCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCCGCTGGGTGCCGTAGCGTGAAGTCTTACGAGCAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCCGCTGGGTGCCGTAGCGTGAAGTCTTACGAGCAAGAG 2460
QY 2461 TTGGCGAAAGATTCACCGTATTTGCCGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATTCACCGTATTTGCCGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAAACATCCGCGAGATGTCGGTCTTCTGTTGATGCGCGTGGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAAACATCCGCGAGATGTCGGTCTTCTGTTGATGCGCGTGGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACAAACAACTAATAAACAACCTCGGATCGAGGAGAGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAACTAATAAACAACCTCGGATCGAGGAGAGATTTTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGAATGTGGTAACTCTTCTCTGAGAGCTCTCACTGCTGCGCCACACAGCGTACCCAC 2820
Db 2761 CCGAATGTGGTAACTCTTCTCTGAGAGCTCTCACTGCTGCGCCACACAGCGTACCCAC 2820
QY 2821 ACSSGTGAACAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACSSGTGAACAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGCTGCGCATCAACCGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGCTGCGCATCAACCGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTCTCTCGTCTGCAAAATCTGTCGCGCAACAAAGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTCTGAGCAGAGCTCCAGCTGTCGCGCAACAAAGTACTCACACCGGGGAGAAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTTGTAAGTCTTCCAGCGCGAGCGATACCTGTTGCGCCAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTTGTAAGTCTTCCAGCGAGAGCTCCCTTGGTGGCGCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTTAGT 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTTAGT 3120
QY 3121 CAGCGCGCCACCTGGCGAGCCATCAACGACTCATCTATGCGGAGAGCCATCAAAATGT 3180
Db 3121 GACTGCGCGACCTGCTCGCCATCAACGACTCATCTATGCGGAGAGCCATCAAAATGT 3180
QY 3181 CCAGAATGTGGCAAGTCTTCTCTCGGTGCAAAATCTGTCGCGGACCAACCGTACTCAC 3240
Db 3181 CCAGAATGTGGCAAGTCTTCTCTCGGTGCAAAATCTGTCGCGGACCAACCGTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGGCGCGCAGTACCCGTCAGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGGCGCGCAGTACCCGTCAGACGTTCCGGACTACGCT 3300
```

RESULT 6

AAH25681
ID AAH25681 standard; DNA; 8101 BP.

XX
AC
XX
DT
XX

05-SEP-2001 (first entry)

XX
Nucleotide sequence of human Cdc25C phosphatase DNA.

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
KW ss.
XX
XX Homo sapiens.
XX
PN WO200144467-A2.
XX
XX 21-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-PR003496.
XX
XX 14-DEC-1999; 95FR-00015722.
PR 30-MAY-2000; 2000FR-00006883.
PR 21-SEP-2000; 2000FR-00012008.
XX
XX (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
PI Goubin-Gramatica F, Ducommun B, Prevost G;
XX
DR WPI; 2001-398152/42.
XX
PT New fusion protein of maltose-binding protein and Cdc phosphatase, useful
PT for identifying phosphatase modulators for regulating the cell cycle.
XX
PS Claim 7; Page 9-13; 56pp; French.
XX
CC The present sequence encodes human Cdc25C phosphatase, and was expressed
CC in Escherichia coli. It is used to make fusion proteins with Escherichia
CC coli maltose binding protein (MBP). The specification describes fusion
CC proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
CC are involved in activation of cyclin-dependent kinases (CDK) that control
CC cell division. The fusion proteins are used to identify modulators of the
CC specified human Cdc phosphatases, potentially useful for regulation of
CC cell division. They are also for studying physiological or
CC physiopathological activities of Cdc phosphatase
XX
SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGATGATAGCGCCGGAAGAGA 60
Db 1 CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGATGATAGCGCCGGAAGAGA 60

Qy 61 GTCATTCAGGGTGTGTAATGTGAACCCAGTAACTGATGATGCGAGATGCGG 120
Db 61 GTCATTCAGGGTGTGTAATGTGAACCCAGTAACTGATGATGCGAGATGCGG 120

Qy 121 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGCCAGCCAGCTTTCTGCGAATA 180
Db 121 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGCCAGCCAGCTTTCTGCGAATA 180

Qy 181 CGCGGAAAGTGAAGCGGATGGCGGAGCTGAATTAATCCCAACCCGCTGGCAC 240
Db 181 CGCGGAAAGTGAAGCGGATGGCGGAGCTGAATTAATCCCAACCCGCTGGCAC 240

Qy 241 AACAACTGGCGGCAACAGTCTGTGATTTGGGCTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCTGTGATTTGGGCTGGCCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGCCCGTTCGCAATTTGCGGGGATTAATCTCGCGCGCATCAACTGGGTGCCAGG 360
Db 301 ACGCCCGTTCGCAATTTGCGGGGATTAATCTCGCGCGCATCAACTGGGTGCCAGG 360

Qy 361 TGGTGGTTCGATGTTAGAACGAGCGGCTGGAAGCCCTGTAAGCGGCGGTGCAATC 420
Db 361 TGGTGGTTCGATGTTAGAACGAGCGGCTGGAAGCCCTGTAAGCGGCGGTGCAATC 420

Qy 421 TTCTCGCGCAACGCGTCACTAGTGGGCTGATCACTAACTATCCGCTGGATGACCAAGGATGCCA 480

Db 421 TTCTCGCGCAACGCGTCACTAGTGGGCTGATCACTAACTATCCGCTGGATGACCAAGGATGCCA 480
Qy 481 TTCTGTGCGAAGCTGCCTGCACTAATGTTCCGGCGGTATTTCTTGATGTTCTCTGACCCAGA 540
Db 481 TTCTGTGCGAAGCTGCCTGCACTAATGTTCCGGCGGTATTTCTTGATGTTCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGGTCAACAGCAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGGCATTTGGGTCAACAGCAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAAACGGAAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAAACGGAAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGCGCATGTTCCCACTGGGATGCTGTTGCCAAGATCAGATGCGGCTGGCGGCA 840
Db 781 ATGAGGCGCATGTTCCCACTGGGATGCTGTTGCCAAGATCAGATGCGGCTGGCGGCA 840
Qy 841 TGCGGCGCATTAACCGAGTCCGGCTGCGGCTGGTGGCGGATATCTCGGTAGTGGGATAGC 900
Db 841 TGCGGCGCATTAACCGAGTCCGGCTGCGGCTGGTGGCGGATATCTCGGTAGTGGGATAGC 900
Qy 901 ACATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAAGGATTTTC 960
Db 901 ACATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGCGCAACCGAGTGGAGCCGCTTGTCTGCAACTCTCTCAGGCGCAGGCGGTA 1020
Db 961 GCCTGCTGGGCGCAACCGAGTGGAGCCGCTTGTCTGCAACTCTCTCAGGCGCAGGCGGTA 1020
Qy 1021 AGGGCAATCAGCTGTTGCGCGCTCTCACTGTTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCGCGCTCTCACTGTTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCAGCAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCAGCAGCAGGTTT 1140
Qy 1141 CCGACTGGAAGCGGCGAGTGAGGCGAACCGCAATTAATGAGTGTAGTCTACTCATTTAG 1200
Db 1141 CCGACTGGAAGCGGCGAGTGAGGCGAACCGCAATTAATGAGTGTAGTCTACTCATTTAG 1200
Qy 1201 GCACAAATTCATGTTTGCACAGCTTATCATCGACTGCGGCTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTCATGTTTGCACAGCTTATCATCGACTGCGGCTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGTGGTATGCGTGTGAGGTTCGTAATCACTGATTAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGCGTGTGAGGTTCGTAATCACTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCGACTCCCGTTCTGGAATTAATGTTTTTGGCGCGCATCATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCGACTCCCGTTCTGGAATTAATGTTTTTGGCGCGCATCATCAACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTATATGTTGTA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTATATGTTGTA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAAGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAAGA 1500
Qy 1501 GCATTTCACCAAGGACCAATAGATTATGAAAACTGAAGAGGTAATCTGGTAATCTGG 1560
Db 1501 GCATTTCACCAAGGACCAATAGATTATGAAAACTGAAGAGGTAATCTGGTAATCTGG 1560

Db 1501 GCACCTTCCACCAAGGACCAGTATATGAAAAATCGAAGAGGTAATACTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATACCGATACCGATACCGATACCGATACCGAT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATACCGATACCGATACCGATACCGATACCGAT 1680
Qy 1681 GCGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCGCTGTTGGCTGAATCAACCGGCAAGCGTTCCAGGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCGCTGTTGGCTGAATCAACCGGCAAGCGTTCCAGGCAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACCAACGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACCAACGCAAGCTGATGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTAACAAAGATCTGCTGCGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTAACAAAGATCTGCTGCGAACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTCGATTAAGAACTGAAGCGAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATTAAGAACTGAAGCGAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAGACGCAAGTACGATTAAGACGTTGCGCTGCGATTAAGCTGGCGGCGAAGCG 2100
Db 2041 TATGAAGACGCAAGTACGATTAAGACGTTGCGCTGCGATTAAGCTGGCGGCGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAACAAACACATGATGATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAACAAACACATGATGATGACACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGTGCCTTTAATAAGCGCAACAGCGATGACCATCAACGCGCGGTGG 2220
Db 2161 TCCATCGCAGAGTGCCTTTAATAAGCGCAACAGCGATGACCATCAACGCGCGGTGG 2220
Qy 2221 GCATGGTCCAACTCGACACGACGAAAGTGAATATGTTGTAACCGTACTGCGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACGACGAAAGTGAATATGTTGTAACCGTACTGCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCGCGCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCGCGCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Qy 2461 TTGCGGAAAGATCCACGTTATGCGGCCCAACATGGAACCGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTTATGCGGCCCAACATGGAACCGCCCAAGAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGCTTCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGACATGTCGATGAAGCGCTGAAAGACGCGCAGACTAATTCAGGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCGCTGAAAGACGCGCAGACTAATTCAGGCTCG 2640

Qy 2641 AACCAACAACATATAACAATAACAACAACTCGGGATCGAGGAGAGATTTTCAGAAATTC 2700
Db 2641 AACCAACAACATATAACAATAACAACAACTCGGGATCGAGGAGAGATTTTCAGAAATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 7
ABLS3238
ID ABL53238 standard; DNA; 8101 BP.
XX ABL53238;
AC ABL53238;
DT 17-JUN-2002 (first entry)
XX Nucleotide sequence of pMAL-Hs Cdc25C.
DE Cdc25C phosphatase; amine; Cdc25 phosphatase; spontaneous alopecia;
KW alopecia; proliferative disease; parasitic disease; viral infection;
KW neurodegenerative disease; myopathy; angiogenesis; psoriasis; restenosis;
KW vitamin K; ss.
XX Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1528..2677
FT /*tag= b
FT /note= "encodes maltose binding protein (MBP)"
FT 2713..4134
FT /*tag= a
FT /note= "ORF of human Cdc25C"
XX FR2812198-A1.
XX
PD 01-FEB-2002.
XX
PF 28-JUL-2000; 2000FR-00009900.
XX
PR 28-JUL-2000; 2000FR-00009900.
XX
PA (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
PI Prevost G, Brezak PMC, Galcera CMO, Thuriereau C, Goubin GF;
PI Ducommun B;
XX
DR WPI; 2002-166066/22.
XX
PT cdc25 Phosphatase inhibiting medicaments, useful e.g. for treating tumor
PT diseases, viral infections, neurodegenerative disease or alopecia,
PT containing new or known aryl-substituted amines.
XX
PS Disclosure; Page 18-22; 39pp; French.
XX
CC The present sequence represents pMAL-HsCdc25C, a plasmid encoding human
CC Cdc25C phosphatase. The specification describes the use of aryl-
CC substituted secondary or tertiary amines for the production of
CC medicaments for inhibiting Cdc25 phosphatases, especially Cdc25C
CC phosphatases. The amines of the invention are used for treating
CC spontaneous alopecia or alopecia induced by exogenous products or
CC radiation. They are also used for treating tumoral or non-tumoral
CC proliferative diseases, parasitic diseases, viral infections,
CC neurodegenerative diseases or myopathy. Examples of non-tumoral
CC proliferative diseases are e.g. angiogenesis, psoriasis or restenosis.
CC The amines are also useful for inhibiting the proliferation of
CC microorganisms (especially yeasts) and in the treatment of all disorders
CC conventionally treated with vitamin K or its derivatives
XX
SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match 82.0%; Score 2705.8; DB 6; Length 8101;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	CGACACCATCAATGCTGCAAAACCTTTTCGGGTATGGCATATAGAGCCCGGAAGAGA	60
Db	1	CGACACCATCGATGGTGCAGAACCTTTTCGGGTATGGCATATAGAGCCCGGAAGAGA	60
Qy	61	GTCAATTCAAGGTGTGAATGTGAACCAAGTAACCTTATACGATGTCCAGAGTATGCG	120
Db	61	GTCAATTCAAGGTGTGAATGTGAACCAAGTAACCTTATACGATGTCCAGAGTATGCG	120
Qy	121	GTGTCCTTATCAGACCTTTTCGGGTGTGAACCAAGCCAGCCAGCCAGCTTTCTCGA	180
Db	121	GTGTCCTTATCAGACCTTTTCGGGTGTGAACCAAGCCAGCCAGCCAGCTTTCTCGA	180
Qy	181	CGCGGAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC	240
Db	181	CGCGGAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC	240
Qy	241	AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Qy	301	AGCGCGTTCGCAAAATTTGTTCGGCGGATTAATCTCGCGCGGATCAATCGGGTGCGAGC	360
Db	301	AGCGCGTTCGCAAAATTTGTTCGGCGGATTAATCTCGCGCGGATCAATCGGGTGCGAGC	360
Qy	361	TGGTGGTTCGATGTAGAACCAAGCGGCTCGAAGCCTGTAAAGCGGCTGCAAACTC	420
Db	361	TGGTGGTTCGATGTAGAACCAAGCGGCTCGAAGCCTGTAAAGCGGCTGCAAACTC	420
Qy	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCCGCTGGATCACCAGGATGCCA	480
Db	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCCGCTGGATCACCAGGATGCCA	480
Qy	481	TTGCTGTGGAAGCTGCCTGCAATAATGTTTCGGCGGTTATTTCTGATGTCTGACACAGA	540
Db	481	TTGCTGTGGAAGCTGCCTGCAATAATGTTTCGGCGGTTATTTCTGATGTCTGACACAGA	540
Qy	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600
Qy	601	TGGTGGCATTTGGTCAACAGCAAACTCGCTGTTAGCGGGCCCAATTAAGTCTGCTCGG	660
Db	601	TGGTGGCATTTGGTCAACAGCAAACTCGCTGTTAGCGGGCCCAATTAAGTCTGCTCGG	660
Qy	661	CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTCAACAAACCATGCAATGCTGA	780
Qy	781	ATGAGGGCATCTTCCCATCGCATGCTGTTTGCACCAACGATCAGATGGCGCTGGGGCAA	840
Db	781	ATGAGGGCATCTTCCCATCGCATGCTGTTTGCACCAACGATCAGATGGCGCTGGGGCAA	840
Qy	841	TGCGGCGCAATACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTGGGATACG	900
Db	841	TGCGGCGCAATACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAAAAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAAAAGGATTTTC	960
Qy	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA	1020
Db	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCCCTGGCGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCCCTGGCGCCCAATA	1080
Qy	1081	CGAAAAACCGCTCTCTCCCGCGGCTTGCAGATTCATTAATGACGACGACGAGTTT	1140
Db	1081	CGAAAAACCGCTCTCTCCCGCGGCTTGCAGATTCATTAATGACGACGACGAGTTT	1140
Qy	1141	CCGACCTGGAAGCGGCGAGTGACGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGACCTGGAAGCGGCGAGTGACGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Qy	1201	GCACAAATTCATGTTGACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGCGG	1260
Db	1201	GCACAAATTCATGTTGACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGCGG	1260
Qy	1261	TCAGGACGCCATCGGAAGCTGTGTATGGCTGTGAGTAAATCACTGCATAAATTCG	1320
Db	1261	TCAGGACGCCATCGGAAGCTGTGTATGGCTGTGAGTAAATCACTGCATAAATTCG	1320
Qy	1321	TGTCGCTCAAGCGGACCTCCCGTTCTGGATTAATGTTTTTTCGGCGGACATCAACGGTT	1380
Db	1321	TGTCGCTCAAGCGGACCTCCCGTTCTGGATTAATGTTTTTTCGGCGGACATCAACGGTT	1380
Qy	1381	CTGGCAATATTTCTGAAATGAGCTTGAACAATTAATCATCGGCTCGTATATGTTGGA	1440
Db	1381	CTGGCAATATTTCTGAAATGAGCTTGAACAATTAATCATCGGCTCGTATATGTTGGA	1440
Qy	1441	ATTGTGAGCGGATAAACAATTTTCACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500
Db	1441	ATTGTGAGCGGATAAACAATTTTCACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500
Qy	1501	GCATTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAACTGG	1560
Db	1501	GCATTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAACTGG	1560
Qy	1561	ATTAAACCGGATAAAGGCTAAACGGTCTCCGCTGAGTCCGTTAGAAATTCGAGAAAGT	1620
Db	1561	ATTAAACCGGATAAAGGCTAAACGGTCTCCGCTGAGTCCGTTAGAAATTCGAGAAAGT	1620
Qy	1621	ACCGGAATTAAGTCAACGGTTCAGCATCCGGATAAACTGGAGAGAAAAATCCACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACGGTTCAGCATCCGGATAAACTGGAGAGAAAAATCCACAGGTT	1680
Qy	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGACACGACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT	1860
Db	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT	1860
Qy	1861	GAAGGTTTATCGCTGATTTATAACAAAGATCTGTGCGGAAACCCGCCAAAAACCTGGGAA	1920
Db	1861	GAAGGTTTATCGCTGATTTATAACAAAGATCTGTGCGGAAACCCGCCAAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAAGGAAAGTGAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAAGGAAAGTGAAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGTGAGCGGGGTTATCGGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGTGAGCGGGGTTATCGGTTCAAG	2040
Qy	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGCGCGAAGCG	2100
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGCGCGAAGCG	2100
Qy	2101	GGTCTGACCTTCTGGTGAACCTGATTTAAAAACAAACATGAATGCGAGACCGGATTC	2160
Db	2101	GGTCTGACCTTCTGGTGAACCTGATTTAAAAACAAACATGAATGCGAGACCGGATTC	2160

QY 601 TGGTCGCATTGGGTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTCGCATTGGGTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGTGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTCGCTGTGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
DB 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTCCCACTGGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGCA 840
DB 781 ATGAGGGCATCGTCCCACTGGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGCA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA 1080
DB 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAGCACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAGCACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTCTACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTCTACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCGGTCACCAATGCTTCTGGCG 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCGGTCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGGTATGCGCTGCGAGTGGTAAATCACTGCGATTAATCG 1320
DB 1261 TCAGCAGCCATCGGAAGCTGGTATGCGCTGCGAGTGGTAAATCACTGCGATTAATCG 1320
QY 1321 TGTGCTCAAGGGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
DB 1321 TGTGCTCAAGGGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
QY 1381 CTGGCAATATTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGTGA 1440
DB 1381 CTGGCAATATTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGTGA 1440
QY 1441 ATTGTAGCGGATAACAATTTACACAGGAACAGCCAGTCCGTTTAGGTGTTTTCAAGA 1500
DB 1441 ATTGTAGCGGATAACAATTTACACAGGAACAGCCAGTCCGTTTAGGTGTTTTCAAGA 1500
QY 1501 GCATTTACCAACAGGACCATAGATTATGAAAACTGAAGAAGTTAACTGTAATCTGG 1560
DB 1501 GCATTTACCAACAGGACCATAGATTATGAAAACTGAAGAAGTTAACTGTAATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTGAAGAAATTCAGAAAGAT 1620
DB 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTGAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAATCGGCGATGGGCCCTGACATTAATCTCTCGGGCACAGACGGCTTTGGTGGCTAC 1740

RESULT 9
ADO23613
ID ADO23613 standard; DNA; 7322 BP.
XX
AC ADO23613;

DB 1681 GCGGCAATCGGCGATGGCCCTGACATTAATCTTCTGGGCACAGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTTGGTGGCTGAAATCAACCCCGGACAAAGCGTTCAGGACAAGCTGAT 1800
DB 1741 GCTCAATCTGGCTTGGTGGCTGAAATCAACCCCGGACAAAGCGTTCAGGACAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTACAACGGCAAGCTGATTGTTTACCCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGTACGTTACAACGGCAAGCTGATTGTTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTCCTGCTGACGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTCCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
DB 2041 TATGAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAATAAACAACAACATGAAATGACAGACCGGATTC 2160
DB 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAATAAACAACAACATGAAATGACAGACCGGATTC 2160
QY 2161 TCCATCGCAGAGTGCCTTTAATAAGGCGAAACAGCGATGACCATCAAGCGGCCGCTGG 2220
DB 2161 TCCATCGCAGAGTGCCTTTAATAAGGCGAAACAGCGATGACCATCAAGCGGCCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACCAAGTAAATATGCTGTAACCGTACTGCGGACCTTC 2280
DB 2221 GCATGCTCAACATCGACACCAAGTAAATATGCTGTAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
DB 2281 AAGGTCACCATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
DB 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGTGCCTGAGCGCTGAAGTCTTACGAGGAAG 2460
DB 2401 GAAGCGTTAATAAGACAAACCGCTGGTGCCTGAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCAAGATCCACGTTATTCGGCCCAACCATGGAACAAACGCCCAAGAGGTGAATCATG 2520
DB 2461 TTGGCAAGATCCACGTTATTCGGCCCAACCATGGAACAAACGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGATGCGGCTGATCAACGCC 2580
DB 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGATGCGGCTGATCAACGCC 2580
QY 2581 GCCAGCGCTGCTGAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
DB 2581 GCCAGCGCTGCTGAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAAACAACCAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
DB 2641 AACCAACAACAATAAACAACCAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

XX	01-JUL-2004	(first entry)
DI	XX	
DE	XX	DNA encoding MBP-ToxoP30del110(52-284aa) fusion protein.
XX	XX	
KW	XX	P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW	XX	MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW	XX	Toxoplasma gondii; db; gene.
XX	XX	
OS	XX	Toxoplasma gondii.
OS	XX	Synthetic.
PN	XX	US2004067239-A1.
XX	XX	
PD	XX	08-APR-2004.
XX	XX	
PF	XX	02-OCT-2002; 2002US-00263153.
XX	XX	
PR	XX	02-OCT-2002; 2002US-00263153.
XX	XX	
PA	XX	(MAIN/) MAINE G T.
PA	XX	(PATE/) PATEL C B.
PA	XX	(GINS/) GINSBURG S R.
PA	XX	(BLIE/) BLIESE T R.
XX	XX	
PI	XX	Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX	XX	
DR	XX	WPI; 2004-304563/28.
DR	XX	P-PSDB; ADO23614.
XX	XX	
PT	XX	Novel purified polypeptide having sequence identity to amino acid
PT	XX	sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT	XX	Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT	XX	antibodies to Toxoplasma gondii.
XX	XX	
PS	XX	Example 2; Fig 25; 114pp; English.
XX	XX	
CC	XX	The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC	XX	defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC	XX	amino acids added to the C-terminus of the amino acid sequence of
CC	XX	Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC	XX	which at least one of the five C-terminal cysteine amino acids of the
CC	XX	amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC	XX	with alanine, or comprising the amino acid sequence chosen from MBP-
CC	XX	Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC	XX	fusion proteins. (I) is useful for detecting the presence of IgM
CC	XX	antibodies to Toxoplasma gondii in a test sample, which involves
CC	XX	contacting the test sample suspected of containing the IgM antibodies
CC	XX	with a composition comprising (I) and detecting the presence of (I)/IgM
CC	XX	antibody complexes. The present sequence represents DNA encoding a MBP-
CC	XX	ToxoP30 fusion protein of the invention.
XX	XX	
SQ	XX	Sequence 7322 BP; 1887 A; 1844 C; 1934 G; 1657 T; 0 U; 0 Other;
XX	XX	
Query Match	81.5%;	Score 2690.4; DB 12; Length 7322;
Best Local Similarity	99.8%;	Pred. No. 0;
Matches 2694;	Conservative	0; Mismatches 6; Indels 0; Gaps 0
Qy	1	CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
Db	1	CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
Qy	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGAGAGTATGCCG 120
Db	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGAGAGTATGCCG 120
Qy	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCAGCTTTTCTCGGAAA 180
Db	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCAGCTTTTCTCGGAAA 180
Qy	181	CGCGGGAAAAGTGGAGCGCGCATGGCGAGCTGAATTACATTCCTCCCAACCGCTGGCAC 240
Db	181	CGCGGGAAAAGTGGAGCGCGCATGGCGAGCTGAATTACATTCCTCCCAACCGCTGGCAC 240

QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCCGACATCATTAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCCGACATCATTAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTA 1440
Db 1381 CTGGCAATATTTCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTA 1440
QY 1441 ATTGTGACGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTCACGA 1500
Db 1441 ATTGTGACGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTCACGA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTATGAAAACAGGAAAGTAAACTGTTATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTATGAAAACAGGAAAGTAAACTGTTATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAAGTGGAAAGAGAAATTCGCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAAGTGGAAAGAGAAATTCGCCACAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGGCTGTTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGGCTGTTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTCCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTCCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTCGCGCGCTGATTCGCTGACGGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTCGCGCGCTGATTCGCTGACGGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGGGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGGGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAACCAACACATCAATGACGACACCGCATTCAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAACCAACACATCAATGACGACACCGCATTCAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAACATCGACACAGCAAGGTAATTTATGGTGAACGTAACGCTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACATCGACACAGCAAGGTAATTTATGGTGAACGTAACGCTGCGGACCTTC 2280
QY 2281 AAGGCTCAACCATCAACCGTTTCTGGCTGCTGAGGCGAGGTAATTAACCGCCCGCAGT 2340
Db 2281 AAGGCTCAACCATCAACCGTTTCTGGCTGCTGAGGCGAGGTAATTAACCGCCCGCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACATCTCTGCTGACTGATCAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACATCTCTGCTGACTGATCAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460

Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTGCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTGCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATTCGCCAGATGTCGCTTTCTGGTATGCGCGTCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATTCGCCAGATGTCGCTTTCTGGTATGCGCGTCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACACAAACAATAACAATAACAACCAACCACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACACACAAACAATAACAATAACAACCAACCACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
RESULT 10
ADO23603
ID ADO23603 standard; DNA; 7352 BP.
XX ADO23603;
XX 01-JUL-2004 (first entry)
XX DNA encoding MBP-ToxoP30del14C(52-294aa) fusion protein.
XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX Toxoplasma gondii.
OS Synthetic.
XX US2004067239-A1.
XX 08-APR-2004.
XX 02-OCT-2002; 2002US-00263153.
XX 02-OCT-2002; 2002US-00263153.
XX (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-FSDB; ADO23604.
XX Novel purified polypeptide having sequence identity to amino acid sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C, Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG antibodies to Toxoplasma gondii.
XX Example 2; Fig 19; 114pp; English.
XX The invention relates to a purified P30 antigen (I) chosen from 3 fully defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6 amino acids added to the C-terminus of the amino acid sequence of Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence of the amino acid sequence of Toxo30del13C P30 antigen sequence is substituted with alanine, or comprising the amino acid sequence chosen from MBP-Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1 fusion proteins. (I) is useful for detecting the presence of IgM antibodies to Toxoplasma gondii in a test sample, which involves contacting the test sample suspected of containing the IgM antibodies

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
XX Toxob30 fusion protein of the invention.

SQ Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGCGCATGATAGCCCGGAGAGA	60
DB	1	CGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGCGCATGATAGCCCGGAGAGA	60
QY	61	GTCAATTACAGGGTGGTGAATGTGAACACAGTAACAGTTATACCATGTTCGCGAGTATGCCG	120
DB	61	GTCAATTACAGGGTGGTGAATGTGAACACAGTAACAGTTATACCATGTTCGCGAGTATGCCG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAA	180
DB	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAA	180
QY	181	CGCGGAAAACTGGAAGCGGATGCGGAGCTGAATTACATTCCTCCAACCGCGTGGCAC	240
DB	181	CGCGGAAAACTGGAAGCGGATGCGGAGCTGAATTACATTCCTCCAACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTGCAGAAATTGTCGGCGGATTAATCTCGCGCGATCAATCGGTGTCACGC	360
DB	301	ACGCGCGTGCAGAAATTGTCGGCGGATTAATCTCGCGCGATCAATCGGTGTCACGC	360
QY	361	TGGTGTGTCGATGTAGAACGAGCGGCTGAAAGCCTGTAAGCGGCTGACCAATC	420
DB	361	TGGTGTGTCGATGTAGAACGAGCGGCTGAAAGCCTGTAAGCGGCTGACCAATC	420
QY	421	TTCTCGCGCAACCGTCACTGAGTGGCTGATCAATTAATCTCGCGCGATCAATCGGATGCCA	480
DB	421	TTCTCGCGCAACCGTCACTGAGTGGCTGATCAATTAATCTCGCGCGATCAATCGGATGCCA	480
QY	481	TTGCTGTGGAAGTCCCTGCATTAATGTTCCGGGTTATTTCTGATGTCCTGACCGA	540
DB	481	TTGCTGTGGAAGTCCCTGCATTAATGTTCCGGGTTATTTCTGATGTCCTGACCGA	540
QY	541	CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGTACGCGTGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGTGGCGTGGAGCATC	600
QY	601	TGGTGGCATTTGGTCCAGCAAAATTCGCGTGTAGCGGGCCCAATTAAGTTCTGTCGG	660
DB	601	TGGTGGCATTTGGTCCAGCAAAATTCGCGTGTAGCGGGCCCAATTAAGTTCTGTCGG	660
QY	661	CGCGTCTGCGTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAAACGGGAAGGACGTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAATGCTGA	780
DB	721	CGGAAACGGGAAGGACGTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGGCATGTTCCCACTCGCATGCTGGTGGCAACGATCAGATGGCGTGGCGCAA	840
DB	781	ATGAGGGCATGTTCCCACTCGCATGCTGGTGGCAACGATCAGATGGCGTGGCGCAA	840
QY	841	TGCGGCCATTACAGGTCCGGGCTGGCGTGTGGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGGCCATTACAGGTCCGGGCTGGCGTGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGTCTATGTTTATTCGCGCGTTAAACCAACATCAAAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGTCTATGTTTATTCGCGCGTTAAACCAACATCAAAACAGGATTTTC	960

QY	961	GCCTGCTGGGGCAAAACAGCGTGCACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGGCAAAACAGCGTGCACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGCCGTCTCAGTGTGAAAGAAAGAAACACCCCTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTGCCGTCTCAGTGTGAAAGAAAGAAACACCCCTGGCGCCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGACAGCTGGCAGCAGAGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGACAGCTGGCAGCAGAGTTT	1140
QY	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG	1200
DB	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG	1200
QY	1201	GCACAAATCTCATGTTTGCAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGCAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGCAGCCATCGAAAGCTGTGATGCTGTGAGTGTGCAATCACTGCATAAATTCG	1320
DB	1261	TCAGCAGCCATCGAAAGCTGTGATGCTGTGAGTGTGCAATCACTGCATAAATTCG	1320
QY	1321	TGTCGCTCAAGCGGCACTCCCGTCTGGATAATGTTTTTTCGCGCGAGCATCAACCGTT	1380
DB	1321	TGTCGCTCAAGCGGCACTCCCGTCTGGATAATGTTTTTTCGCGCGAGCATCAACCGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTGTGTGACAAATTAATCATCGCTGTGATAATGTTGGA	1440
DB	1381	CTGCAAAATTTCTGAAATGAGCTGTGTGACAAATTAATCATCGCTGTGATAATGTTGGA	1440
QY	1441	ATTGTGAGCGGATTAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTCACGA	1500
DB	1441	ATTGTGAGCGGATTAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTCACGA	1500
QY	1501	GCATTCACCAACAAAGGACCATAGATTATGAAACTGAAAGGTAACCTGTTAACTCTGG	1560
DB	1501	GCATTCACCAACAAAGGACCATAGATTATGAAACTGAAAGGTAACCTGTTAACTCTGG	1560
QY	1561	ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACCAAGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACCAAGTT	1680
QY	1681	CGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGCTTGGTGGCTAC	1740
DB	1681	CGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGCTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGAAACAAAGCGTTCCAGGACAAAGCTGTAT	1800
DB	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGAAACAAAGCGTTCCAGGACAAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATCGCGTACGTTACAGCGCAAGCTGATTTGCTTACCGGATCGCTGTT	1860
DB	1801	CCGTTTACCTGGGATCGCGTACGTTACAGCGCAAGCTGATTTGCTTACCGGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAAACCCGCAAAACCTCGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAAACCCGCAAAACCTCGGAA	1920
QY	1921	GAGATCCCGCGCTGATAAGAACTGAAAGGTAAGGTAAGAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGATAAGAACTGAAAGGTAAGGTAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG	2040

Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGCGACTGGCGGTGAGCATC 600
QY 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CCGCTCTCGCTCTGGCTGGCTGGCTGACATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGCTCTCGCTCTGGCTGGCTGACATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAAGCGGAAGCGCACTGGAGTGCATGTCCGGTTTCAACAAACCAATGCAATGCTGA 780
Db 721 CCGAAGCGGAAGCGCACTGGAGTGCATGTCCGGTTTCAACAAACCAATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGCGATGTGGTTGCCAAGATCAGATGCGCTGGCGCGCA 840
Db 781 ATGAGGGCATCGTTCCCACTGCGATGTGGTTGCCAAGATCAGATGCGCTGGCGCGCA 840
QY 841 TCGCGGCCAATACAGATCCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCCAATACAGATCCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTATTATCCCGCGCTTAACCAACCAATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATTATCCCGCGCTTAACCAACCAATCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTGCGGCCAATA 1080
QY 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTATGCGAGCTGCGACGAGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTATGCGAGCTGCGACGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTATGTTGACAGCTTATCATCGACTGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1260
Db 1201 GCACAATTTCTATGTTGACAGCTTATCATCGACTGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCGTGTGCGAGTCAATCACTGCAATAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCGTGTGCGAGTCAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCAGCATATAACCGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCAGCATATAACCGTT 1380
QY 1381 CTGGCAATATTTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAATATTTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTCAACAGGAACAGCAGTCCGTTTATAGTGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCAACAGGAACAGCAGTCCGTTTATAGTGTGTTTTCACGA 1500
QY 1501 GCATTTCAACCAAGGACCATAGATATGAAAACTGAAGAGGTAAATCTGGTAATCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATATGAAAACTGAAGAGGTAAATCTGGTAATCTGG 1560
QY 1561 ATTAACGCGATAAGGCTATAACGGTCTCGCTGAGTTCGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGCGATAAGGCTATAACGGTCTCGCTGAGTTCGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAGTT 1680

QY 1681 CGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTTGTTTACCOCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTTGTTTACCOCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCAACCOCGCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCAACCOCGCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCTGATTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCTGATTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAAGACGTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAAGACGTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAACAACATGATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAACAACATGATGACACACCGATTAC 2160
QY 2161 TCCATCCGAGAGCTGCTTTAATAAGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTAATAAGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
QY 2221 GCATGTCCAAACATCGACACCAAGTGAATTTATGGTGTAAACGCTGCTGCGACCTTC 2280
Db 2221 GCATGTCCAAACATCGACACCAAGTGAATTTATGGTGTAAACGCTGCTGCGACCTTC 2280
QY 2281 AAGGTCACCACTCCAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCACTCCAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCNACAAAGAGCTGGCAAAAGATTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCNACAAAGAGCTGGCAAAAGATTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTTGCCGCAACATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTTGCCGCAACATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCNACATCCGAGATGCTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCNACATCCGAGATGCTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTCGTGAGACTGTGAGTGAAGCCCTGAAAGAGCGCGAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCGTGAGACTGTGAGTGAAGCCCTGAAAGAGCGCGAGACTTAATTCGAGCTCG 2640
QY 2641 AACACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700

RESULT 12

AD023639

ID AD023639 standard; DNA; 7370 BP.

XX	ADO23639;
XX	AC
XX	DT 01-JUL-2004 (first entry)
XX	DNA encoding MBP-Toxop30MIX1 fusion protein.
XX	DE
XX	P30 antigen; Toxo30del3C; Toxo30del2C; Toxop30 MIX1;
KW	MBP-Toxo30del3C(52-300aa); MBP-Toxo30del4C(52-294aa); MBP-Toxo30MIX1;
KW	Toxoplasma gondii; ds; gene.
XX	OS
XX	Toxoplasma gondii.
OS	Synthetic.
XX	XX
PN	US2004067239--A1.
XX	XX
PD	08-APR-2004.
XX	XX
PF	02-OCT-2002; 2002US-00263153.
XX	XX
XX	02-OCT-2002; 2002US-00263153.
XX	(MAIN//) MAINE G T.
PA	(PATE//) PATEL C B.
PA	(GINS//) GINSBURG S R.
XX	(BLIE//) BLIESE T R.
XX	XX
PI	Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX	XX
DR	WFI; 2004-304563/28.
XX	P-PSDB; ADO23640.
XX	Novel purified polypeptide having sequence identity to amino acid sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del3C, Toxo30del2C, Toxop30 MIX1, useful for detecting presence of IgM or IgG antibodies to Toxoplasma gondii. Example 5; Fig 31; 114pp; English.
XX	XX
CC	The invention relates to a purified p30 antigen (I) chosen from 3 fully defined Toxo30del3C, Toxo30del2C and Toxop30 MIX1 sequences, having 1-6 amino acids added to the C-terminus of the amino acid sequence of Toxo30del2C p30 antigen sequence, or comprising an amino acid sequence in which at least one of the five C-terminal cysteine amino acids of the amino acid sequence of Toxo30del3C p30 antigen sequence is substituted with alanine, or comprising the amino acid sequence chosen from MBP- Toxo30del3C(52-300aa), MBP-Toxo30del4C(52-294aa) and MBP-Toxo30MIX1 fusion proteins. (I) is useful for detecting the presence of IgM antibodies to Toxoplasma gondii in a test sample, which involves contacting the test sample suspected of containing the IgM antibodies with a composition comprising (I) and detecting the presence of (I)/IgM antibody complexes. The present sequence represents DNA encoding a MBP- Toxop30 fusion protein of the invention.
XX	Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;
XX	Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
XX	Best Local Similarity 99.8%; Pred No. 0;
XX	Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0
Qy	1 CCGCACCATCGAATGGTGCAAAACCTTTTCGCCGTATGCATGATAGCGCCCGGAAGA 60
Db	1 CCGCACCATCGAATGGTGCAAAACCTTTTCGCCGTATGCATGATAGCGCCCGGAAGA 60
Qy	61 GTCATTCCAGGTTGTAATGTGAAACCAGTAACTGTTATACGATGTCGACAGTAGTATCGCG 120
Db	61 GTCATTCCAGGTTGTAATGTGAAACCAGTAACTGTTATACGATGTCGACAGTAGTATCGCG 120
Qy	121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGACCCAGCACGTTTCTGCGAAAA 180
Db	121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGACCCAGCACGTTTCTGCGAAAA 180
Qy	181 CGCGGGAAAAAGTGGAAACGCGCGATGTGGCGGAGCTGAAATTTACATTTCCACACCGCTGGCAC 240

181	CGCGGAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTTACATTTCCAAACCGCGTGGCAC	240
241	AACAACTGGCGGGCAAAACAGTCGTTCCTGCTGATTTGGCGTTGGCACCTCAGTCTTGGCCCTGCG	300
241	AACAACTGGCGGGCAAAACAGTCGTTCCTGCTGATTTGGCGTTGGCACCTCAGTCTTGGCCCTGCG	300
301	ACGCGCGTTCGCAAAATTTGTCGGGCGGATTAATCTCGCGCGGATCAACTCTGGGTGCCAGCG	360
301	ACGCGCGTTCGCAAAATTTGTCGGGCGGATTAATCTCGCGCGGATCAACTCTGGGTGCCAGCG	360
361	TGCTGTGTTCGATGTGTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGCGGTGCCAATC	420
361	TGCTGTGTTCGATGTGTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGCGGTGCCAATC	420
421	TTCTCTCGCAACCGCGTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGATGCCA	480
421	TTCTCTCGCAACCGCGTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGATGCCA	480
481	TTGCTGTGAAGGTGCGCTGCGACTAAATGTTTCGGCGGTTATTTCTTGATGTCCTTGACGAGA	540
481	TTGCTGTGAAGGTGCGCTGCGACTAAATGTTTCGGCGGTTATTTCTTGATGTCCTTGACGAGA	540
541	CACCCATCAACAGTATATTTCTCCCATCAAGACAGGTAGCGCACTGGGCGTGGAGCATC	600
541	CACCCATCAACAGTATATTTCTCCCATCAAGACAGGTAGCGCACTGGGCGTGGAGCATC	600
601	TGCTGCGATTGGGTCAACAGCAATCCGCTGTGTTAGCGGCGCCATTAAGTTCTGTCTCGG	660
601	TGCTGCGATTGGGTCAACAGCAATCCGCTGTGTTAGCGGCGCCATTAAGTTCTGTCTCGG	660
661	CGCGTCTGCGTCTGGCTGGTGCATATAATCTCACTCGCAATCAAAATCAGCCGATAG	720
661	CGCGTCTGCGTCTGGCTGGTGCATATAATCTCACTCGCAATCAAAATCAGCCGATAG	720
721	CGGAAACGGNAGCGACTGGAGTGCATATGTCGCCGTTTTCGAACAAACATGCAATGCTGA	780
721	CGGAAACGGNAGCGCGACTGGAGTGCATATGTCGCCGTTTTCGAACAAACATGCAATGCTGA	780
781	ATGAGGCGATCGTTTCCACATCGCATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCAA	840
781	ATGAGGCGATCGTTTCCACATCGCATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCAA	840
841	TGGCGCGCATTAACGAGTCCGGCTCGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900
841	TGGCGCGCATTAACGAGTCCGGCTCGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900
901	ACGATACCGAAGACAGCTCATGTTTATATCCGCGGTTTAAACCATCAAAACAGGATTTTC	960
901	ACGATACCGAAGACAGCTCATGTTTATATCCGCGGTTTAAACCATCAAAACAGGATTTTC	960
961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
1021	AGGGCAATCAGCTGTTGGCCCGTCTCACTGGTGAAGAAAAACACCTCGCGCCCAATA	1080
1021	AGGGCAATCAGCTGTTGGCCCGTCTCACTGGTGAAGAAAAACACCTCGCGCCCAATA	1080
1081	CGCAAAACCGGCTCTCCCGCGGCTGGCCGATTCATTAATGCACTGGCAGCAGGTTT	1140
1081	CGCAAAACCGGCTCTCCCGCGGCTGGCCGATTCATTAATGCACTGGCAGCAGGTTT	1140
1141	CCGACCTGGAAGCGGCGACTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
1141	CCGACCTGGAAGCGGCGACTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
1201	GCACAAATCTCATGTTTGAAGAGCTTATCATCTGACTGACGCGTGACCAATGCTTCTGGCG	1260
1201	GCACAAATCTCATGTTTGAAGAGCTTATCATCTGACTGACGCGTGACCAATGCTTCTGGCG	1260
1261	TCAGGCAAGCCATCGGAAGCTGTGGTATGGCTGTGACGCTCGTAATCACTGCATAATTCG	1320

Db 1261 TCAGGAGGCCATCGAAGCTGTGGTATGCTGTGCAGGTCTGTAATCACTGCTAATTCG 1320
QY 1321 TGTGCGTCAAGGGGCACTCCGCTTCCTGGATAATGTTTTTGGCCGACATCAACGGTT 1380
Db 1321 TGTGCGTCAAGGGGCACTCCGCTTCCTGGATAATGTTTTTGGCCGACATCAACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATAATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCGGTTAGTGTTTTCAAGA 1500
Db 1441 ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCGGTTAGTGTTTTCAAGA 1500
QY 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAAACTGAAAGAGGTAAACTGCTAATCTGG 1560
Db 1501 GCACCTTCAACCAACAGGACCATAGCATATGAAATTCGAAGAGGTAAACTGCTAATCTGG 1560
QY 1561 ATTAAACGGGATAAAGGCTATTAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATAAAGGCTATTAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGAGTCAACCGTTCAGCATCCGATTAACCTGGAAGAGGTAAACTGCTAATCTGG 1680
Db 1621 ACCGGAATTAAGAGTCAACCGTTCAGCATCCGATTAACCTGGAAGAGGTAAACTGCTAATCTGG 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGACAAAGCGTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGACAAAGCGTTCAGGACAAGCTGTAT 1800
QY 1801 CGGTTTACTCGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCGCATCGCTGTT 1860
Db 1801 CGGTTTACTCGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCGCATCGCTGTT 1860
QY 1861 GAAGGTTATCGCTGATTTAATAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTAATAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGATAAGAACTGAAGGAAAGTAAAGAGGAGGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGATAAGAACTGAAGGAAAGTAAAGAGGAGGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGAAGCTGGGCGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGAAGCTGGGCGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGCTGCTGATTAATAACAACACATCAATGATGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGCTGCTGATTAATAACAACACATCAATGATGACACCGGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCACATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGCACTTC 2280
Db 2221 GCATGGTCCACATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGCACTTC 2280
QY 2281 AAGGGTCAACCATCAACACCGCTTCTGTTGGCTGCTGAGCGCAGGTATTAACGGCCGAGT 2340
Db 2281 AAGGGTCAACCATCAACACCGCTTCTGTTGGCTGCTGAGCGCAGGTATTAACGGCCGAGT 2340
QY 2341 CCGAACAAAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400

QY 2401 GAAGCGGTTAATAAGACAAACCCCTGGTGGTCCCTAGCGTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCCCTGGTGGTCCCTAGCGTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGCGGAAAGATCCACGTTATTTGGCCGACCAATGCAAAACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTTATTTGGCCGACCAATGCAAAACGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGATATGCCGTGGCTACTCGGGTGTATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGATATGCCGTGGCTACTCGGGTGTATCAACGCC 2580
QY 2581 GCCAGCGTCTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
QY 2641 AACCAACAAACAATAAACAATAACAACAACCTCGGGATCGAGGAAGGATTCAGAAATTC 2700
Db 2641 AACCAACAAACAATAAACAATAACAACAACCTCGGGATCGAGGAAGGATTCAGAAATTC 2700

RESULT 13
ADO23649
ID ADO23649 standard; DNA; 7370 BP.
XX ADO23649;
AC ADO23649;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-ToxoP30MIX5 fusion protein.
XX
KW P30 antigen; Toxo30del3C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del3C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
XX WPI; 2004-304563/28.
XX P-PSDB; ADO23650.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del3C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
XX Example 5; Fig 35; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del3C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence of
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del12C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del12C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM

CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the Igm antibodies
CC with a composition comprising (i) and detecting the presence of (i)/Igm
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATATAGGCCCGGAGAGA 60
DB 1 CGGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATATAGGCCCGGAGAGA 60
QY 61 GTCNAATTCAGGGTGGTGAATGTGAACACAGTAACCTTATACGATGTCGAGAGTATGCG 120
DB 61 GTCNAATTCAGGGTGGTGAATGTGAACACAGTAACCTTATACGATGTCGAGAGTATGCG 120
QY 121 GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCCAGCCAGTTTCTGCGAAAA 180
DB 121 GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCCAGCCAGTTTCTGCGAAAA 180
QY 181 CGCGGGAAGATGGAACGGCGATGGCGAGCTGAATTAACATTCCTCAACCGCGTGGCAC 240
DB 181 CGCGGGAAGATGGAACGGCGATGGCGAGCTGAATTAACATTCCTCAACCGCGTGGCAC 240
QY 241 AACAACTCGCGGGCAACAGTCTGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
DB 241 AACAACTCGCGGGCAACAGTCTGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGGCCCTGCGCAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
DB 301 ACGGCCCTGCGCAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCTGTGAAGCGGCGGTGCACAATC 420
DB 361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCTGTGAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACGAGGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACGAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTCCCTGCAATAATGTTCCGGCGTTATTTCTTGATGTTCTGTACCCAGA 540
DB 481 TTGCTGTGGAAGCTCCCTGCAATAATGTTCCGGCGTTATTTCTTGATGTTCTGTACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGCAATGGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
DB 601 TGGTCGCAATGGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTCGCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTCGCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGTGGCGGCA 840
DB 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGTGGCGGCA 840
QY 841 TGCAGCCATTAACAGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB 841 TGCAGCCATTAACAGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTTATATATCCCGCGCTTAACCAACCATCAAAAGGATTTTC 960

DB 901 ACGATACCGAAGACAGCTCATGTTTATATATCCCGCGTTAAACCAACATCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAAAACACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCTCTCACTGTGTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCTCTCACTGTGTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCAGTGGCAGCAGAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCAGTGGCAGCAGAGGTTT 1140
QY 1141 CCCGACTGGAAGGCGGCGAGTGCAGCGCAACCGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGGCGGCGAGTGCAGCGCAACCGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATTCGTTCTGGCG 1260
DB 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATTCGTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATAAATTCG 1320
DB 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTTGATTAATGTTTTGCGCGGACATCATACCGGTT 1380
DB 1321 TGTGCTCAAGCGCACTCCCGTCTTGATTAATGTTTTGCGCGGACATCATACCGGTT 1380
QY 1381 CTGCAAAATTTCTGAAATGAGTGTGCAAAATTAATCATCGGCTCGTAAATCACTGCATAAATTCG 1440
DB 1381 CTGCAAAATTTCTGAAATGAGTGTGCAAAATTAATCATCGGCTCGTAAATCACTGCATAAATTCG 1440
QY 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAACACCGGCTCGGTTAGGTTGTTTCAAGA 1500
DB 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAACACCGGCTCGGTTAGGTTGTTTCAAGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAACTGAAGAAAGGTAATACTGGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAACTGAAGAAAGGTAATACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGGAAATTCAGAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGGAAATTCAGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAACTGGAAGAGAAATTCACACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAACTGGAAGAGAAATTCACACAGGTT 1680
QY 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
DB 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
DB 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGGCGTACGTTAACAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGGCGTACGTTAACAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040

Db 1981 CTCAAGAACCGTACTTCACTGCGCGTGAATTGCTGCTGACGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGGCTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGGCTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTTGAACCTGATTAAGAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTTGAACCTGATTAAGAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACGACGACGAGTGAATTTATGGTGAACGCTACTCGCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACGACGAGTGAATTTATGGTGAACGCTACTCGCGACCTTC 2280
QY 2281 AAGGTCACACCTCCAAACCGTTCGTTGGCTGCTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCTCCAAACCGTTCGTTGGCTGCTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGCGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CCGAACAAAGAGCTGCGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAACTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAACTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCAGGTATTCGCGCCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGGTATTCGCGCCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGACATGTCGCTGCTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCGCGACATGTCGCTGCTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTCTGTCAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACACACACATTAACATTAACAAACCTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
Db 2641 AACACACACACATTAACATTAACAAACCTCGGGATCGAGGGAAGATTTCAGAAATTC 2700

RESULT 14
ADO23644
XX ADO23644 standard; DNA; 7370 BP.
AC ADO23644;
XX ADO23644;
DT 01-JUL-2004 (first entry)
DE DNA encoding MBP-ToxoP30MIX3 fusion protein.
KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
FN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.

PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
DR WPI; 2004-304563/28.
DR P-FSDB; ADO23645.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of Igm or Igm
PT antibodies to Toxoplasma gondii.
XX
XX Example 5; Fig 33; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of Igm
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the Igm antibodies
CC with a composition comprising (I) and detecting the presence of (I)/Igm
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATCGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATCGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCAATTCAGGTTGGTGAATGTGAAACCAAGTAACTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGTTGGTGAATGTGAAACCAAGTAACTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGGCGACAGCTTTCTGCGAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGGCGACAGCTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTGTGCTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTGTGCTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTGTAAGCAAGCGCGCTGGAAGCCCTGTAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTGATGTGTAAGCAAGCGCGCTGGAAGCCCTGTAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCTGCACTTAATTTCCGGCGTTATTTCTGATGTTCTTGACCA 540
Db 481 TTGCTGTGGAAGCTGCTGCTGCACTTAATTTCCGGCGTTATTTCTGATGTTCTTGACCA 540

QY 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGGCACTGGCGTGGAGCATC 600
QY 601 TGGTCGCATTGGGTCAACAGAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTCAACAGAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGGCTGGCTGGCGTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTGGCTGGCGTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGACCGGAAGGCACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGACCGGAAGGCACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGTGTCACACGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGTGTCACACGATCAGATGGCGCTGGCGCGAA 840
QY 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTATTATCCCGCCGTTAACACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATTATCCCGCCGTTAACACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAATCCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAATCCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTGAGTTAGTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTGAGTTAGTCACTCATTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATGACCTGCACGCTGCACCAATGCTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATGACCTGCACGCTGCACCAATGCTCTGGCG 1260
QY 1261 TCAGGACGCATCGGAAGCTGTGGTATGGCTGTGACAGTCTGTAATCACTGATTAATTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGTGGTATGGCTGTGACAGTCTGTAATCACTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTACGA 1500
QY 1501 GCACCTTACCACAGGACCATAGCATATGAAATTCGAAGAAGGTAACTGGTAATCTGG 1560
Db 1501 GCACCTTACCACAGGACCATAGCATATGAAATTCGAAGAAGGTAACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAAAGAGAAATTCACACAGGTT 1680

Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAACTGGAAAGAGAAATTCACACAGGTT 1680
QY 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCGGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCGGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGSTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGSTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTAAGAACCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGAACCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAATGATGAGACACCGATTAAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAATGATGAGACACCGATTAAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGTTCAACATCGACACCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGTTCAACATCGACACCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTTCAACCATCAAAACCGTTGTTGGCTGTGAGCGCAGGTATTAACCGCCCACT 2340
Db 2281 AAGGTTCAACCATCAAAACCGTTGTTGGCTGTGAGCGCAGGTATTAACCGCCCACT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCAATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCAATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700

QY	181	CGCGGAAAAGTGGAAAGCGGGGATGGCGGAGCTGAATTACATTCCTCCAAACCGCTGGCAC	240
DB	181	CGCGGAAAAGTGGAAAGCGGGGATGGCGGAGCTGAATTACATTCCTCCAAACCGCTGGCAC	240
QY	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGGCCACTTCAGTCTGGGCCCTGC	300
DB	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGGCCACTTCAGTCTGGGCCCTGC	300
QY	301	ACGCGCGTTCGGAATTTGTTCGCGGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
DB	301	ACGCGCGTTCGGAATTTGTTCGCGGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
QY	361	TGTTGGTCTCGATGTTGTAAGAAAGAGCGGCTGCGAGCTGTAAAGCGGCGTGCACAATC	420
DB	361	TGTTGGTCTCGATGTTGTAAGAAAGAGCGGCGTGCAGGCTGTAAAGCGGCGTGCACAATC	420
QY	421	TTCTCGCGCAACGGCTCAGTGGGCTGATCACTAACTATCCGCTGGATGACACAGGATGCCA	480
DB	421	TTCTCGCGCAACGGCTCAGTGGGCTGATCACTAACTATCCGCTGGATGACACAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCCTGCCTAAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540
DB	481	TTGCTGTGGAAGCTGCCTGCCTAAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540
QY	541	CACCCATCAAACAGTATTAATTTCTCCCATGAAGACGGTACCGCATCGGCGTGGAGCATC	600
DB	541	CACCCATCAAACAGTATTAATTTCTCCCATGAAGACGGTACCGCATCGGCGTGGAGCATC	600
QY	601	TGTTGCGCATTTGGGTACACAGCAAAATCGGCTGTTAGCGGGGCCATTAAGTTCTGTCGG	660
DB	601	TGTTGCGCATTTGGGTACACAGCAAAATCGGCTGTTAGCGGGGCCATTAAGTTCTGTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAAACGGGAAGGCGACTGAGTGCCATGTCCGTTTTTCAACAAACCATGCAATGCTGA	780
DB	721	CGGAAACGGGAAGGCGACTGAGTGCCATGTCCGTTTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCGATCTGTTCCCACTTCGATGCTGGTTGCGCAACGATCAGATGCGCTGGCGGCA	840
DB	781	ATGAGGCGATCTGTTCCCACTTCGATGCTGGTTGCGCAACGATCAGATGCGCTGGCGGCA	840
QY	841	TGCGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTATGCGGATACG	900
DB	841	TGCGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTATGCGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGTTAAACCAATCAAAACAGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGTTAAACCAATCAAAACAGATTTTC	960
QY	961	GCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
DB	961	GCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGCCGCTCTACTGTGTAAAGAAAAACCACTGGCGGCCAATA	1080
DB	1021	AGGGCAATCAGCTGTTGCCGCTCTACTGTGTAAAGAAAAACCACTGGCGGCCAATA	1080
QY	1081	CGAAACCGGCTCTCCCGCGCGTTGGCCGATTCACTAAATGCACTGGCACGACAGGTTT	1140
DB	1081	CGAAACCGGCTCTCCCGCGCGTTGGCCGATTCACTAAATGCACTGGCACGACAGGTTT	1140
QY	1141	CCGCACTGGAAAGCGGCGAGTGAAGCAACGCAATTAATGTGATTTAGTCACTCATTAG	1200
DB	1141	CCGCACTGGAAAGCGGCGAGTGAAGCAACGCAATTAATGTGATTTAGTCACTCATTAG	1200
QY	1201	GCACAATTTCTATGTTTGAAGCTTATCATCGACTGCAAGGTCACCAATGCTTCTGGCG	1260
DB	1201	GCACAATTTCTATGTTTGAAGCTTATCATCGACTGCAAGGTCACCAATGCTTCTGGCG	1260

Qy	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATTAATTCG	1320
Db	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATTAATTCG	1320
Qy	1321	TGTGCTCAAGCGGCACTCCCGTTCGTGATAATGTTTTTGGCGGACATCATAAACGGTT	1380
Db	1321	TGTGCTCAAGCGGCACTCCCGTTCGTGATAATGTTTTTGGCGGACATCATAAACGGTT	1380
Qy	1381	CTGGCAATATTTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTGA	1440
Db	1381	CTGGCAATATTTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTGA	1440
Qy	1441	ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500
Qy	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGGAAGAGTAAACTGTTAATCTGG	1560
Db	1501	GCACCTTCAACCAAGGACCATAGCATATGAAATTCGAAGAGGTAATCTGGTAATCTGG	1560
Qy	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAAATTCGAGAAAGAT	1620
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAAATTCGCCAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAAATTCGCCAGGTT	1680
Qy	1681	CGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCGGACAAAGCGTTCAGGACAAGCTGAT	1800
Db	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCGGACAAAGCGTTCAGGACAAGCTGAT	1800
Qy	1801	CCGTTTACCTGGGATGCCGTACGTTTACAAACGGCAAGCTGATTCCTACCCGATCGCTGT	1860
Db	1801	CCGTTTACCTGGGATGCCGTACGTTTACAAACGGCAAGCTGATTCCTACCCGATCGCTGT	1860
Qy	1861	GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGGTACTTCACTGGCGGCTGATGCTGTGACGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGGTACTTCACTGGCGGCTGATGCTGTGACGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAARAACGCAAGTACGACATTAAGACGTGGCGGTGGATAACGCTGGCGGAAAGCG	2100
Db	2041	TATGAARAACGCAAGTACGACATTAAGACGTGGCGGTGGATAACGCTGGCGGAAAGCG	2100
Qy	2101	GCTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC	2160
Db	2101	GCTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC	2160
Qy	2161	TCCATCGAGAAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGAGAAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Qy	2221	GCATGGTCCACATCGACACGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGGTCCACATCGACACGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGTTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT	2340
Db	2281	AAGGTTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT	2340
Qy	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTTCCTCGAAAACTATCTGCTGACTGATGAAGGTC	2400

Search completed: August 1, 2005, 04:28:30
Job time : 1104.6 secs

Db	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTC	2400
Qy	2401	GAAGCGGTTAATAAAGACAAACCCGCTGGGTCCGCTAGCGCTGAAGTCTTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCCGCTGGGTCCGCTAGCGCTGAAGTCTTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAAAGATCCACGTTATTGCGCCGCAACATGGAAAAACGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAAGATCCACGTTATTGCGCCGCAACATGGAAAAACGCCAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG	2640
Qy	2641	AACAAACAAACAATAACAATAACAACAACTCCGGATCGAGGGAAGGATTTTCAGAATTC	2700
Db	2641	AACAAACAAACAATAACAATAACAACAACTCCGGATCGAGGGAAGGATTTTCAGAATTC	2700

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgcaccatcgatgtgc.....acgacgttcgcgactaagct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2700	81.8	7475	US-08-971-036-1	Sequence 1, Appli
2	2700	81.8	7475	US-09-096-570-1	Sequence 1, Appli
3	2700	81.8	7475	US-09-265-617B-1	Sequence 1, Appli
c 4	1417.6	43.0	5926	US-09-027-169-3	Sequence 3, Appli
5	1362	41.3	5201	US-09-640-882-2	Sequence 2, Appli
6	1362	41.3	5201	US-09-640-882-3	Sequence 3, Appli
7	1204	36.5	4557	US-08-778-717-5	Sequence 5, Appli
8	1200.8	36.4	3832	US-08-148-675A-2	Sequence 2, Appli
c 9	1199.2	36.3	5248	US-08-487-283A-18	Sequence 18, Appli
c 10	1199.2	36.3	5248	PT-US96-05611A-21	Sequence 21, Appli
c 11	1199.2	36.3	5312	US-10-263-103-35	Sequence 35, Appli
c 12	1199.2	36.3	5443	US-08-929-967-1	Sequence 1, Appli
c 13	1199.2	36.3	5502	US-09-702-705-785	Sequence 785, App
c 14	1199.2	36.3	5502	US-09-736-457-785	Sequence 785, App
c 15	1199.2	36.3	5502	US-09-614-124B-785	Sequence 785, App
c 16	1199.2	36.3	5502	US-09-671-325-785	Sequence 785, App
c 17	1199.2	36.3	5502	US-09-589-184-785	Sequence 785, App
c 18	1199.2	36.3	5502	US-09-658-824-785	Sequence 785, App
c 19	1199.2	36.3	5616	US-08-929-967-3	Sequence 3, Appli
c 20	1199.2	36.3	5873	US-09-695-437A-62	Sequence 62, Appli
c 21	1199.2	36.3	6353	US-09-702-705-784	Sequence 784, App
c 22	1199.2	36.3	6353	US-09-736-457-784	Sequence 784, App
c 23	1199.2	36.3	6353	US-09-614-124B-784	Sequence 784, App
c 24	1199.2	36.3	6353	US-09-671-325-784	Sequence 784, App
c 25	1199.2	36.3	6353	US-09-589-184-784	Sequence 784, App
c 26	1199.2	36.3	6353	US-09-658-824-784	Sequence 784, App
27	1199.2	36.3	6361	US-08-646-538-7	Sequence 7, Appli

c 28	1199.2	36.3	6361	3	US-09-503-222-7	Sequence 7, Appli
c 29	1199.2	36.3	6727	2	US-08-125-462-2	Sequence 2, Appli
c 30	1199.2	36.3	6727	2	US-08-891-848-2	Sequence 2, Appli
c 31	1199.2	36.3	6799	2	US-08-125-462-5	Sequence 5, Appli
c 32	1199.2	36.3	6799	2	US-08-891-848-5	Sequence 5, Appli
c 33	1199.2	36.3	7659	3	US-09-128-314-4	Sequence 4, Appli
c 34	1199.2	36.3	7676	3	US-09-056-556-213	Sequence 213, App
c 35	1199.2	36.3	7676	3	US-09-072-596-208	Sequence 208, App
c 36	1199.2	36.3	7676	4	US-09-072-967-213	Sequence 213, App
c 37	1199.2	36.3	7676	4	US-09-287-849-9	Sequence 9, Appli
c 38	1199.2	36.3	8031	3	US-09-643-597-254	Sequence 254, App
c 39	1199.2	36.3	8031	4	US-09-480-884A-254	Sequence 254, App
c 40	1199.2	36.3	8031	4	US-09-542-615A-254	Sequence 254, App
c 41	1199.2	36.3	8031	4	US-09-606-421B-254	Sequence 254, App
c 42	1199.2	36.3	8031	4	US-09-476-496A-254	Sequence 254, App
c 43	1199.2	36.3	8031	4	US-09-630-940B-254	Sequence 254, App
c 44	1199.2	36.3	8157	3	US-09-128-314-3	Sequence 3, Appli
c 45	1199.2	36.3	8501	3	US-08-793-900-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; LENGTH: 7475 base pairs
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8% Score 2700; DB 2; Length 7475;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CGGACACCATGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
Db 1 CGGACACCATGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
QY 61 GTCAATTGAGGTGGTGAATGTGAACACCAAGTAACTGATATACGATGTCGACAGTATGCCG 120
Db 61 GTCAATTGAGGTGGTGAATGTGAACACCAAGTAACTGATATACGATGTCGACAGTATGCCG 120
QY 121 GTGTCCTCTATCAGACCGTTTCCCGCGTGGTGAACACCGGCAGGCACGCTTTCTCGAAAA 180
Db 121 GTGTCCTCTATCAGACCGTTTCCCGCGTGGTGAACACCGGCAGGCACGCTTTCTCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCATGGCGGAGCTGAATTAATATCCCAACCGCTGCGAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTAATATCCCAACCGCTGCGAC 240
QY 241 AACAACTGGCGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCTCGCAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGGCGCTCGCAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGGTGAACGAGCGCGTGAAGCGCTGAAAGCGCGGTGACAAATC 420
Db 361 TGGTGGTTCGATGGTGAACGAGCGCGTGAAGCGCTGAAAGCGCGGTGACAAATC 420
QY 421 TTCTGGCGCAACGCGTCAGTGGGCTGATCAATTAATATCCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTGGCGCAACGCGTCAGTGGGCTGATCAATTAATATCCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGGTTATTTCTTGATGTTCTCTGACGAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGGTTATTTCTTGATGTTCTCTGACGAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGACTGGCGGTGAGGATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGACTGGCGGTGAGGATC 600
QY 601 TGGTGCATTTGGGTGCAACAGCAATTCGCGTGTGTAGCGGGCCCAATTAAGTTCTGTCGG 660
Db 601 TGGTGCATTTGGGTGCAACAGCAATTCGCGTGTGTAGCGGGCCCAATTAAGTTCTGTCGG 660
QY 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCGATGTTCCCACTCGCATGCTGTTGTCGCAACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGCGATGTTCCCACTCGCATGCTGTTGTCGCAACGATCAGATGGCGTGGCGCAA 840
QY 841 TGC CGCATTTACCGAGTTCGGGCTGCGGCTGGTGGTGGATATCTCGGTAGTGGGATACG 900
Db 841 TGC CGCATTTACCGAGTTCGGGCTGCGGCTGGTGGTGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTAAACCCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTAAACCCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGAACGCTTCTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGAACGCTTCTGCAACTCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGCAATACAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATACAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
```

```
Db 1021 AGGCAATACAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGAAAAACCGCTCTTCCCGCGCGTGGCGGATTCATTAATGCAAGTGGCAGCAGAGTTT 1140
Db 1081 CGAAAAACCGCTCTTCCCGCGCGTGGCGGATTCATTAATGCAAGTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGACGCGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGACGCGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGACCATCGGAAGCTGTGATGGCTGTGACGAGTGTAAATCACTGCAATAATTCG 1320
Db 1261 TCAGGACGACCATCGGAAGCTGTGATGGCTGTGACGAGTGTAAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCCGTTCGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTCAACAGGAAAAACGCGAGTCCGTTTATGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTCAACAGGAAAAACGCGAGTCCGTTTATGTTTTCACGA 1500
QY 1501 GCACCTTCAACCAACAAAGGACCATAGATTATGAAAACTGGAAGAGTAACTGGTAACTCGG 1560
Db 1501 GCACCTTCAACCAACAAAGGACCATAGATTATGAAAACTGGAAGAGTAACTGGTAACTCGG 1560
QY 1561 ATTAACGGCGATTAAGGCTATAACCGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATTAAGGCTATAACCGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAGTT 1680
QY 1681 CGCGCAACTGGCGCTGCGCTGACATTTCTTCTGGGCGACACGACCGCTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGGCGCTGCGCTGACATTTCTTCTGGGCGACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCAAGCTGATTCCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCAAGCTGATTCCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAGAGCGGCTGTATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAGAGCGGCTGTATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGACGGGGGTTATGTTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGACGGGGGTTATGTTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTAAGAACTGAAAGCGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGAACTGAAAGCGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTCAGCTTCTGCTGTTGACCTGATTAAACACACATCAATGAATGACACCGGATTAC 2160
Db 2101 GGTCTCAGCTTCTGCTGTTGACCTGATTAAACACACATCAATGAATGACACCGGATTAC 2160
```

QY 2161 TCCATCGCAGAGTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
DB 2161 TCCATCGCAGAGTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCACATCGACACAGCAAGTGAATTAATGTTAAACGGTACTCGCGACCTTC 2280
DB 2221 GCATGGTCCACATCGACACAGCAAGTGAATTAATGTTAAACGGTACTCGCGACCTTC 2280
QY 2281 AAGGTCACCATCAACACCGTTGCTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
DB 2281 AAGGTCACCATCAACACCGTTGCTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
QY 2341 CGAACAAGAGCTGGCAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
DB 2341 CGAACAAGAGCTGGCAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
DB 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTCGCCGCCACCATGGAAACGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGGCGAAAGATCCACGATTTCGCCGCCACCATGGAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CGAACAATCCCGCAGATGTCGCTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
DB 2521 CGAACAATCCCGCAGATGTCGCTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGACACTGTCGATGAAGCCCTGAAAGACGCCGACACTAATTCGAGCTCG 2640
DB 2581 GCCAGCGTGTGACACTGTCGATGAAGCCCTGAAAGACGCCGACACTAATTCGAGCTCG 2640
QY 2641 AACACAACAAACAATAACAATAACAACCTCGGGATCGAGGAGGATTTCAAGATTC 2700
DB 2641 AACACAACAAACAATAACAATAACAACCTCGGGATCGAGGAGGATTTCAAGATTC 2700

RESULT 2
US-09-096-570-1
; Sequence 1, Application US/09096570
; Patent No. 6018020
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/096, 570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreibler, Lewis J
; REGISTRATION NUMBER: 38522

; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-09-096-570-1

Query Match 81.8%; Score 2700; DB 3; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACCATCAATGTTGCAAAACCTTTCGGGTATGGCATGATAGCGCCGGAGAGA 60
DB 1 CCGACACCATCAATGTTGCAAAACCTTTCGGGTATGGCATGATAGCGCCGGAGAGA 60
QY 61 GTCAATTCCAGGTGTTGAATGTAAGAACCCAGTAAAGTTATACGATCTCGCAGAGTATGCG 120
DB 61 GTCAATTCCAGGTGTTGAATGTAAGAACCCAGTAAAGTTATACGATCTCGCAGAGTATGCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
DB 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTATATCCCAACCGCGTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTATATCCCAACCGCGTGGCAC 240
QY 241 AACAACTCGCGGGCAACAGTCGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
DB 241 AACAACTCGCGGGCAACAGTCGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTGTGCGCGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
DB 301 ACGCGCGTTCGCAAAATTGTGCGCGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGTTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
DB 361 TGGTGGTTCGATGTTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGTGATGACACGAGGATGCCA 480
DB 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGTGATGACACGAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGGGCTGGGCTGGAGCATC 600
DB 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGGGCTGGGCTGGAGCATC 600
QY 601 TGGTCCGATTGGGTCACCGCAAAATTCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTCCGATTGGGTCACCGCAAAATTCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGTGGCGGCA 840
DB 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGTGGCGGCA 840

QY 841 TCGCGCCATTACCGAGTCGGGCTGCGGTTGGTCGGATATCTCGTAGTGGATACG 900
DB 841 TCGCGCCATTACCGAGTCGGGCTGCGGTTGGTCGGATATCTCGTAGTGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACCGAGTGGACCGCTTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAACACCGAGTGGACCGCTTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCAGCTGGCACACAGGTTT 1140
DB 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCAGCTGGCACACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGAGTTAGTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGATGCTGTGAGAGTGCAGTCACTGATTAATTCG 1320
DB 1261 TCAGGACGCAATCGGAAGCTGTGATGCTGTGAGAGTGCAGTCACTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTCTCGGATTAATGTTTTCGCGGACATCAACCGTT 1380
DB 1321 TGTGCTCAAGCGGCACTCCCGTCTCGGATTAATGTTTTCGCGGACATCAACCGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
DB 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTAGCGGATTAACAAATTTTACAGGAACAGCCAGTCCGTTAGGTGTTTTCAGGA 1500
DB 1441 ATTGTAGCGGATTAACAAATTTTACAGGAACAGCCAGTCCGTTAGGTGTTTTCAGGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAGTAAACTGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAGTAAACTGTAATCTGG 1560
QY 1561 ATTAAACGCGATTAAGGCTATAACGCTCTCGCTGAACTGCGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAAACGCGATTAAGGCTATAACGCTCTCGCTGAACTGCGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATCCCAAGGTT 1680
DB 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATCCCAAGGTT 1680
QY 1681 CGCGCAACTGCGGATGCGGCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
DB 1681 CGCGCAACTGCGGATGCGGCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGCAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGCAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGCTGATTAACCGGCAAGCTGTTGCTTACCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGGCTGATTAACCGGCAAGCTGTTGCTTACCGATCGCTGTT 1860
QY 1861 GAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
DB 1861 GAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920

QY 1921 GAGATCCGCGCTGGATTAAGAACTGAAACGGAAGGTAAAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCGCGCTGGATTAAGAACTGAAACGGAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTCCTGCTGACGGGGTATGCGGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTCCTGCTGACGGGGTATGCGGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
DB 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAGAAACCAACACATGATGACACCGGATTAC 2160
DB 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAGAAACCAACACATGATGACACCGGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCGCGTGG 2220
DB 2161 TCCATCGCAAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCGCGTGG 2220
QY 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGGACCTTC 2280
DB 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGCTCAACCATCCAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACGGCGCGCAGT 2340
DB 2281 AAGGCTCAACCATCCAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACGGCGCGCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGAAGGTTCTG 2400
DB 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGAAGGTTCTG 2400
QY 2401 GAAGCGGTTTAAAGAACCAACCGCTGGGTGCGTAGCGCTGAAGTCTTAACGAGGAAGAG 2460
DB 2401 GAAGCGGTTTAAAGAACCAACCGCTGGGTGCGTAGCGCTGAAGTCTTAACGAGGAAGAG 2460
QY 2461 TTGGCAAAAGATCCACGATTTGCGCCACCATCGAAAAACCGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGGCAAAAGATCCACGATTTGCGCCACCATCGAAAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAAACATCCGAGATGTCGCGCTTTTCTGGTATGCGCTGCTGCTGCGTGAACGCC 2580
DB 2521 CCGAAACATCCGAGATGTCGCGCTTTTCTGGTATGCGCTGCTGCTGCGTGAACGCC 2580
QY 2581 GCAGCGGCTGTCAGACTGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG 2640
DB 2581 GCAGCGGCTGTCAGACTGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
DB 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700

RESULT 3

US-09-265-617B-1
; Sequence 1, Application US/09265617B
; Patent No. 6372883
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R.
; APPLICANT: Hurst, David N.
; APPLICANT: Jones, Philip S.
; APPLICANT: Kay, Paul B.
; APPLICANT: Raynham, Tony M.
; APPLICANT: Wilson, Francis X.
; TITLE OF INVENTION: Antiviral Medicaments
; FILE REFERENCE: 20052 antiviral medicaments
; CURRENT APPLICATION NUMBER: US/09/265,617B
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: GB9806815.8
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

; LENGTH: 7475
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; INFORMATION: Description of Artificial Sequence: Vector and
; OTHER INFORMATION: Gene Fragments
US-09-265-617B-1

Query Match 81.8%; Score 2700; DB 3; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGACACATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGAAGAGA 60

Qy 61 GTCAATTACAGGTGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 120
Db 61 GTCAATTACAGGTGTGAATGTGAACACAGTATACGATGTCGACAGTATGCCG 120

Qy 121 GTGCTCTTATCAGACCCGTTTCCCGCGTGGTGAACACGCGCACCGTTCCTGCGAAAA 180
Db 121 GTGCTCTTATCAGACCCGTTTCCCGCGTGGTGAACACGCGCACCGTTCCTGCGAAAA 180

Qy 181 CCGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAAATTAACAATCCCAACCGCGTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAAATTAACAATCCCAACCGCGTGGCAC 240

Qy 241 AACAACTCGCGGCAAAACAGTGTGCTGATTGGCGTGGCCACTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTCGCGGCAAAACAGTGTGCTGATTGGCGTGGCCACTCCAGTCTGGGCCCTGC 300

Qy 301 AGCGCCGTTCGCAAAATGTTCGCGGCGATTAATCTTCGCGCGCATCAACTGGGTGCCAGG 360
Db 301 AGCGCCGTTCGCAAAATGTTCGCGGCGATTAATCTTCGCGCGCATCAACTGGGTGCCAGG 360

Qy 361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCCGTGAAGCGCGGTGCAATC 420
Db 361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCCGTGAAGCGCGGTGCAATC 420

Qy 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480

Qy 481 TTGCTGTGGAAGCTGCTGCAATATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCAATATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540

Qy 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

Qy 601 TGGTTCGATGGTTCACGCAAAATCGGCTGTAGCGGCGCCATTAAGTTCTGCTCGG 660
Db 601 TGGTTCGATGGTTCACGCAAAATCGGCTGTAGCGGCGCCATTAAGTTCTGCTCGG 660

Qy 661 CCGCTCTGGCTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGCTCTGGCTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Qy 721 CCGAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780

Qy 781 ATGAGGGCATCGTTCCCATGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGGGCAA 840
Db 781 ATGAGGGCATCGTTCCCATGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGGGCAA 840

Qy 841 TCGCGGCAATTAACGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATCG 900
Db 841 TCGCGGCAATTAACGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATCG 900

Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCCATCAAAACAGGATTTTC 960

Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAACAGGATTTTC 960

Qy 961 GCCTGCTGGGGCAAAACAGCGGTGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACAGCGGTGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020

Qy 1021 AGGGCAATCAGCTGTGCCCCGTCTCACTGGTGAAGAAACCAACCTCGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCCCCGTCTCACTGGTGAAGAAACCAACCTCGCGGCCAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGCGACAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGCGACAGAGTTT 1140

Qy 1141 CCCGACTTGGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTATTAG 1200
Db 1141 CCCGACTTGGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTATTAG 1200

Qy 1201 GCACAAATTCATGTTTGAAGCTTATCATCGACTGCAAGCTGCAACCAATGCTTCGCGG 1260
Db 1201 GCACAAATTCATGTTTGAAGCTTATCATCGACTGCAAGCTGCAACCAATGCTTCGCGG 1260

Qy 1261 TCAGCAGCCATCGAAGCTGTGATGCTGCTGAGTGTGAGCTGCTGAGTGTGCTGCTGCTG 1320
Db 1261 TCAGCAGCCATCGAAGCTGTGATGCTGCTGAGTGTGAGCTGCTGAGTGTGCTGCTGCTG 1320

Qy 1321 TGTGCTCAAGCGCGCACTCCCGTCTTGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCGCACTCCCGTCTTGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380

Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATCACTGCTG 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATCACTGCTG 1440

Qy 1441 ATTGTAGCGGATTAACAAATTTTCAACAGGAAACAGCGAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATTAACAAATTTTCAACAGGAAACAGCGAGTCCGTTTAGTGTTCACGA 1500

Qy 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAGTGAAGGTAAGTCTGCTGCTGCTGCTG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAGTGAAGGTAAGTCTGCTGCTGCTGCTG 1560

Qy 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCAGGTT 1680

Qy 1681 GCGGCACTGCGGATGGCGCTGACATTAATCTTGGGCAACAGCCGCTTGGTGGCTTAC 1740
Db 1681 GCGGCACTGCGGATGGCGCTGACATTAATCTTGGGCAACAGCCGCTTGGTGGCTTAC 1740

Qy 1741 GCTCAATTCGCTGTTGGCTGAAATCAACCGGCAAAAGGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATTCGCTGTTGGCTGAAATCAACCGGCAAAAGGTTCCAGGACAAAGCTGAT 1800

Qy 1801 CCGTTTACCTGGGATGCGGTAGCTTAAACCGCAAGCTGATTCCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTAAACCGCAAGCTGATTCCTTACCCGATCGCTGTT 1860

Qy 1861 GAAGCGTTATCGCTGATTTAAACAAAGATCTGCTGCGAAACCCGCCAACCCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTAAACAAAGATCTGCTGCGAAACCCGCCAACCCTGGGAA 1920

Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC 1980

Qy 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGAGCGGGGTTATCGGTTCAAG 2040

Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCGTGTATGCTGTGACGGGGTATGCTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGCTGGATACGCTGGCGGCAAGACG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGCTGGATACGCTGGCGGCAAGACG 2100
QY 2101 GGTCTGACCTTCTGCTGGTGAACCTGATTAATAAACAACACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGGTGAACCTGATTAATAAACAACACATGAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTAAATAAGGCGGCAAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTAAATAAGGCGGCAAGCGATGACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATATGCTGTAACGCTGCTGCCAGCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATATGCTGTAACGCTGCTGCCAGCTTC 2280
QY 2281 AAGGTCACCATCAACACCGTTCGTTGCGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCATCAACACCGTTCGTTGCGCTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGGCTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGGAAAGATCCACGATTTGCGGCCACCAATGGAAGAACGCGCCAGAAAGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGATTTGCGGCCACCAATGGAAGAACGCGCCAGAAAGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGCTCGCTTCTGCTGATGCGGTGCTGCTGCTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGCTCGCTTCTGCTGATGCGGTGCTGCTGCTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
QY 2641 AACAAACAACAATTAACAATTAACAACCTCGGGATCGAGGGAAGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATTAACAATTAACAACCTCGGGATCGAGGGAAGATTCAGAAATTC 2700

RESULT 4

US-09-027-169-3/c
; Sequence 3, Application US/09027169
; Patent No. 6420524
; GENERAL INFORMATION:
; APPLICANT: CRAIG, NANCY L
; TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
; TITLE OF INVENTION: ATP-DEPENDENT TRANSCRIPTION PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Anne Brown (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave.
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027.169
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Brown, Anne
; REGISTRATION NUMBER: 36,463
; REFERENCE/DOCKET NUMBER: 5789-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2205
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "pEM delta R.adj to 1"
; US-09-027-169-3

Query Match 43.0%; Score 1417.6; DB 3; Length 5926;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 9; Indels 12; Gaps 2;
QY 3 GACACCATCGAATGGTGCMAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGT 62
Db 1564 GACACCATCGAATGGTGCMAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGT 1505
QY 63 CAATTCAAGGTGGTGAATGTGAACACCAAGTAACTATACGATGTCGAGAGTATGCCGGT 122
Db 1504 CAATTCAAGGTGGTGAATGTGAACACCAAGTAACTATACGATGTCGAGAGTATGCCGGT 1445
QY 123 GTCTTTATCAGACCGTTTCCCGGTGGTGAACACAGGCCAGCCACGTTTCTGCGAAACG 182
Db 1444 GTCTTTATCAGACCGTTTCCCGGTGGTGAACACAGGCCAGCCACGTTTCTGCGAAACG 1385
QY 183 CGGGAAGAGTGAAGCGCGGATGCGGAGTGAATTTACATTTCCCAACCGCGTGGCAAA 242
Db 1384 CGGGAAGAGTGAAGCGCGGATGCGGAGTGAATTTACATTTCCCAACCGCGTGGCAAA 1325
QY 243 CAATCGCGGCGCAACAGTCTGCTGATTTGGGTTGCCACCTCCAGTCTGGGCCCTGCAC 302
Db 1324 CAATCGCGGCGCAACAGTCTGCTGATTTGGGTTGCCACCTCCAGTCTGGGCCCTGCAC 1265
QY 303 GCGCGTGGCAAAATTTGTCGGCGGATTAATTTCTCGCGCCGATCAACTGGGTGCCAGGTG 362
Db 1264 GCGCGTGGCAAAATTTGTCGGCGGATTAATTTCTCGCGCCGATCAACTGGGTGCCAGGTG 1205
QY 363 GTGCTGCGATGTAGAAACGAGCGGCTGAGAGCTGTAAGCGGGTGCACAACTTT 422
Db 1204 GTGCTGCGATGTAGAAACGAGCGGCTGAGAGCTGTAAGCGGGTGCACAACTTT 1145
QY 423 CTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGGATGCCATT 482
Db 1144 CTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGGATGCCATT 1085
QY 483 GCTGTGGAAGCTGCTGCACTAATTTCTCCCAAGAGACGCTACGCGACTGGGCGTGGAGCATCTG 542
Db 1084 GCTGTGGAAGCTGCTGCACTAATTTCTCCCAAGAGACGCTACGCGACTGGGCGTGGAGCATCTG 1025
QY 543 CCCATCAACAGTATTAATTTCTCCCAAGAGACGCTACGCGACTGGGCGTGGAGCATCTG 602
Db 1024 CCCATCAACAGTATTAATTTCTCCCAAGAGACGCTACGCGACTGGGCGTGGAGCATCTG 965
QY 603 GTCGATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTTAAGTTCTGTCTCGCG 662
Db 964 GTCGATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTTAAGTTCTGTCTCGCG 905
QY 663 CGTCTCGCTGCTGGCTGGCATTAATTTCTCACTCGGAATCAAAATTCAGCGGATAGCG 722
Db 904 CGTCTCGCTGCTGGCTGGCATTAATTTCTCACTCGGAATCAAAATTCAGCGGATAGCG 845
QY 723 GAACGGGAAGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGAAT 782
Db 844 GAACGGGAAGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGAAT 785
QY 783 GAGGGCATCTTCCCACTGCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGGGCGCAATG 842

Db 784 GAGGCACTGTTCCACATGCGATGCTGGTGGCAACGATCAGATGCGCTGGCGCAATG 725
Qy 843 CGCGCATTAACCGAGTCGCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGATACGAC 902
Db 724 CGCGCATTAACCGAGTCGCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGATACGAC 665
Qy 903 GATACCGAAGACAGCTCATGTTATATCCGCGGTAAACCAACCATCAACAGATTTTCG 962
Db 664 GATACCGAAGACAGCTCATGTTATATCCGCGGTCAACCAACCATCAACAGATTTTCG 605
Qy 963 CTGCTGGGGCAAAACAGCGTGGACCGTCTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 604 CTGCTGGGGCAAAACAGCGTGGACCGTCTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 545
Qy 1023 GGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCAACCCCTGGCGCCCAATACG 1082
Db 544 GGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCAACCCCTGGCGCCCAATACG 485
Qy 1083 CAACCGGCTCTCCCGCGCGTGGCCGATTCATTAATGCGAGCTGGCAGCAGGTTTCC 1142
Db 484 CAACCGGCTCTCCCGCGCGTGGCCGATTCATTAATGCGAGCTGGCAGCAGGTTTCC 425
Qy 1143 CGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGAGTTAGCTCACTATTAGGC 1202
Db 424 CGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGAGTTAGC-----GC 376
Qy 1203 ACAATTTCTCATGTTTGACAGCTTATCATCGACTGCACCGTGCACCAATGCTTCTGGCGTC 1262
Db 375 GAATTTGATCTGGTTTGACAGCTTATCATCGACTGCACCGTGCACCAATGCTTCTGGCGTC 316
Qy 1263 AGGAGCCATCGGAAGCTGTGATGCTGTGCGAGTGTGCAAGTGTAAATCACTGCATAATTCGTG 1322
Db 315 AGGAGCCATCGGAAGCTGTGATGCTGTGCGAGTGTGCAAGTGTAAATCACTGCATAATTCGTG 256
Qy 1323 TCGCTCAAGGGCGCACTCCGCTTCTCGATAAGTGTGCGCGGACATCAACCGTTCT 1382
Db 255 TCGCTCAAGGGCGCACTCCGCTTCTCGATAAGTGTGCGCGGACATCAACCGTTCT 196
Qy 1383 GGCAATATTTCTGAATGAGCTGTTGACAATTAATCAT- CGGCTCGTATAATGTTGGAA 1441
Db 195 GGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCCGCTCGTATAATGTTGGAA 136
Qy 1442 TTGTAGCGGATAACAATTTTACACAGGAACAGCC 1477
Db 135 TTGTAGCGGATAACAATTTTACACAGGAACAGAC 100

RESULT 5

US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE2
US-09-640-882-2

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
Qy 3 GACACCATCGAATGCTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCGAAGAGAGT 62
Db 866 GACACCATCGAATGCTGCGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCGAAGAGAGT 925
Qy 63 CAATTCAGGGTGGTGAATGTGAACACCAAGTAAAGTATACGATGTCGAGAGATATGCCGT 122
Db 926 CAATTCAGGGTGGTGAATGTGAACACCAAGTAAAGTATACGATGTCGAGAGATATGCCGT 985
Qy 123 GTCTCTTATCAGACCGTTTCCGCGTGGTGAACAGCGCCAGCCAGCTTTCTGCGAARACG 182
Db 986 GTCTCTTATCAGACCGTTTCCGCGTGGTGAACAGCGCCAGCCAGCTTTCTGCGAARACG 1045
Qy 183 CGGGAAGGTTGAAGCGGCGATGGCGAGCTGAATTAATCTCCCAACCGGTCGACAA 242
Db 1046 CGGGAAGGTTGAAGCGGCGATGGCGAGCTGAATTAATCTCCCAACCGGTCGACAA 1105
Qy 243 CAATGGCGGCAACAGTCTGATGGCGTGGCCACTCCAGTCTGGCCCTGCAC 302
Db 1106 CAATGGCGGCAACAGTCTGATGGCGTGGCCACTCCAGTCTGGCCCTGCAC 1165
Qy 303 CGCGCTGCGAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTCGACGCGTG 362
Db 1166 CGCGCTGCGAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTCGACGCGTG 1225
Qy 363 GTGGTGTGATGATGAACAGCGGCTGCAAGCTCTGAAGCGGCGGTGCACAACTCT 422
Db 1226 GTGGTGTGATGATGAACAGCGGCTGCAAGCTCTGAAGCGGCGGTGCACAACTCT 1285
Qy 423 CTCGCGCAACCGTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACAGGATGCCATT 482
Db 1286 CTCGCGCAACCGTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACAGGATGCCATT 1345
Qy 483 CTTGTGGAGTCTGCTGCACTTAATGTTCCGCGGTATTTCTTGATGTCCTGACAGACA 542
Db 1346 CTTGTGGAGTCTGCTGCACTTAATGTTCCGCGGTATTTCTTGATGTCCTGACAGACA 1405
Qy 543 CCCATCAACAGTATTTATTTCTCCATGAAGACGTCGCGACTGGGCGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTTATTTCTCCATGAAGACGTCGCGACTGGGCGTGGAGCATCTG 1465
Qy 603 CTCGCAATGGGTCAACAGCAAAATCGCGTGTAGCGGCCCAATTAAGTCTGTCTCGGCG 662
Db 1466 CTCGCAATGGGTCAACAGCAAAATCGCGTGTAGCGGCCCAATTAAGTCTGTCTCGGCG 1525
Qy 663 CGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 1526 CGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAACGGGAAGCGACTGGAGTGCATGTCGCGTTCACAAACCCATGCAAAATGCTGAAT 782
Db 1586 GAACGGGAAGCGACTGGAGTGCATGTCGCGTTCACAAACCCATGCAAAATGCTGAAT 1645
Qy 783 GAGGSCATCGTTCCCATCTGCGATGCTGTTGTCACCAACGATCAGATGGCGCTGGGCGCAATG 842
Db 1646 GAGGSCATCGTTCCCATCTGCGATGCTGTTGTCACCAACGATCAGATGGCGCTGGGCGCAATG 1705
Qy 843 CGCGCATTAACCGAGTCTGGCGCTGGCGGATGATCTCGGTAGTGGATACGAC 902
Db 1706 CGCGCATTAACCGAGTCTGGCGCTGGCGGATGATCTCGGTAGTGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGATTTTCG 1825
Qy 963 CTGCTGGGGCAAAACAGCGTGGACCGTGTGTCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 1826 CTGCTGGGGCAAAACAGCGTGGACCGTGTGTCAACTCTCTCAGGGCCAGCGGTGAAG 1885
Qy 1023 GGCAATCAGCTGTTGCCCGTCTCACTGTGTAAGAAAAACCAACCCCTGGCGCCCAATACG 1082

Db 1886 GGCATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACACCTGCGCCCAATAGC 1945
Qy 1083 CAAACCGCTCTCCCGCGGTTGGCCGATTCAATATGACGTGGCAGCAGAGTTTCC 1142
Db 1946 CAAACCGCTCTCCCGCGGTTGGCCGATTCAATATGACGTGGCAGCAGAGTTTCC 2005
Qy 1143 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTATTAGGC 1202
Db 2006 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAG--CGCAATTGATC 2063
Qy 1203 ACAATTCTCATGTTTGAAGCTTATCATCGACTGCAOAGTGCAOAGTCTTCTGGCGTC 1262
Db 2064 TGAATTCTCATGTTTGAAGCTTATCATCGACTGCAOAGTGCAOAGTCTTCTGGCGTC 2123
Qy 1263 AGCAGCCATCGAAGCTGTGTGTATGGCTGTGCAGTCTGTAATCACTGCATTAATTCGTG 1322
Db 2124 AGCAGCCATCGAAGCTGTGTGTATGGCTGTGCAGTCTGTAATCACTGCATTAATTCGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACGTTTCT 1382
Db 2184 TCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACGTTTCT 2243
Qy 1383 GGCATAATTCT 1394
Db 2244 GGCATAATTCT 2255

RESULT 6

US-09-640-882-3
; Sequence 3, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
US-09-640-882-3

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 3 GACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGCCGCCGGAAGAGAGT 62
Db 866 GACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGCCGCCGGAAGAGAGT 925
Qy 63 CAATTTCAGGGTGGTAATGTGAACACCGATACGTTATACGATTCGAGAGTATGCCGGT 122
Db 926 CAATTTCAGGGTGGTAATGTGAACACCGATACGTTATACGATTCGAGAGTATGCCGGT 985
Qy 123 GTCTCTTATCAGACGTTTCCCGGTGTGTGAACACCGATACGTTTCTGCCAAAACG 182
Db 986 GTCTCTTATCAGACGTTTCCCGGTGTGTGAACACCGATACGTTTCTGCCAAAACG 1045
Qy 183 CGGAAAAAGTGGAAAGCGGCGATCGCGAGTGTGAATTAATTCCTCCAAACCGCGTGCAAA 242
Db 1046 CGGAAAAAGTGGAAAGCGGCGATCGCGAGTGTGAATTAATTCCTCCAAACCGCGTGCAAA 1105

Qy 243 CAACTGGCGGCAAAACAGTGGTTCCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGGAC 302
Db 1106 CAACTGGCGGCAAAACAGTGGTTCCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGGAC 1165
Qy 303 GCGCGCTCGCAAAATTTGTCGGCGGCGATTAAATCTCGCCCGGATCAACTGGGTGCCAGGCTG 362
Db 1166 GCGCGCTCGCAAAATTTGTCGGCGGCGATTAAATCTCGCCCGGATCAACTGGGTGCCAGGCTG 1225
Qy 363 GTGTGTGATGTTAGTAAAGCGGCGTGAAGCGCTGTAAAGCGGCGGTGCAAACTCTT 422
Db 1226 GTGTGTGATGTTAGTAAAGCGGCGTGAAGCGGCTGTAAAGCGGCGGTGCAAACTCTT 1285
Qy 423 CTCGCGCAACGCGTCAGTGGCGTATCAATTAATCTCGCTGGATGACCGAGTGCATTT 482
Db 1286 CTCGCGCAACGCGTCAGTGGCGTATCAATTAATCTCGCTGGATGACCGAGTGCATTT 1345
Qy 483 GCTGTGAAAGCTGCTGCACCTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACGACACA 542
Db 1346 GCTGTGAAAGCTGCTGCACCTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACGACACA 1405
Qy 543 CCCATCAACAGTATTTTCTCCATGAAGACGTTACGGCATCTGGCGGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTTTCTCCATGAAGACGTTACGGCATCTGGCGGTGGAGCATCTG 1465
Qy 603 GTCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTTAGTTCTGTCTCGGCG 662
Db 1466 GTCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTTAGTTCTGTCTCGGCG 1525
Qy 663 GGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 722
Db 1526 GGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 1585
Qy 723 GAACGGAAAGCGACCTGGAGTGCATCTCCGTTTTCACAAACCATGCAAAATGCTGAAT 782
Db 1586 GAACGGAAAGCGACCTGGAGTGCATCTCCGTTTTCACAAACCATGCAAAATGCTGAAT 1645
Qy 783 GAGGCGATCTGCTCCACTGCGATGCTGTTGCAACAGATCAGATGGCGTGGCGGCAATG 842
Db 1646 GAGGCGATCTGCTCCACTGCGATGCTGTTGCCAACGATCAGATGGCGTGGCGGCAATG 1705
Qy 843 GCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAGCAG 902
Db 1706 GCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAGCAG 1765
Qy 903 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTCG 1825
Qy 963 CTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGAAG 1022
Db 1826 CTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGAAG 1885
Qy 1023 GGCATACAGCTGTGGCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATAGC 1082
Db 1886 GGCATACAGCTGTGGCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATAGC 1945
Qy 1083 CAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGAGCTGTCAGCAGCAGAGTTTCC 1142
Db 1946 CAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGAGCTGTCAGCAGCAGAGTTTCC 2005
Qy 1143 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTATTAGGC 1202
Db 2006 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAG--CGCAATTGATC 2063
Qy 1203 ACAATTCTCATGTTTGAAGCTTATCATCGACTGCAOAGTGCAOAGTCTTCTGGCGTC 1262
Db 2064 TGAATTCTCATGTTTGAAGCTTATCATCGACTGCAOAGTGCAOAGTCTTCTGGCGTC 2123
Qy 1263 AGCAGCCATCGAAGCTGTGTGTATGGCTGTGCAGTCTGTAATCACTGCATTAATTCGTG 1322
Db 2124 AGCAGCCATCGAAGCTGTGTGTATGGCTGTGCAGTCTGTAATCACTGCATTAATTCGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACGTTTCT 1382

Db 2184 TCGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCCCGACATCAATACGGTTCT 2243

Qy 1383 GCGAAATATCT 1394

Db 2244 GCGAAATATCT 2255

RESULT 7
US-08-778-717-5
; Sequence 5, Application US/08778717
; Patent No. 6602689
; GENERAL INFORMATION:
; APPLICANT: UENO, EIICHI
; APPLICANT: NOBUYUKI, FUJII
; APPLICANT: OKADA, MASAHISA
; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN
; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR
; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/778,717
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 352225/1995
; FILING DATE: 28-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2084-031-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4557 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; ORIGINAL SOURCE:
; ORGANISM: E. COLI
; STRAIN: BL21 (DE3)
; PUBLICATION INFORMATION:
; AUTHORS: NOBUYUKI FUJII ET AL,
; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED
; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF
; TITLE: EXPRESSING SAID FUSED PROTEIN
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557

US-08-778-717-5
Query Match 36.5%; Score 1204; DB 4; Length 4557;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCGACACCATCAATGGTGCAGAACTTTCGGCGGTATGCGATGATAGCGCCGGAAGAGA 60
Db 2563 CCGACACCATCAATGGTGCAGAACTTTCGGCGGTATGCGATGATAGCGCCGGAAGAGA 2622

Qy 61 GTCAATTCAAGGTGGTGAATGTGAAACCAAGTAAAGTTATACGATGTCGACAGTATGCCG 120
Db 2623 GTCAATTCAAGGTGGTGAATGTGAAACCAAGTAAAGTTATACGATGTCGACAGTATGCCG 2682
Qy 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCAGCCAGTCTTCGCAAAA 180
Db 2683 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCAGCCAGTCTTCGCAAAA 2742
Qy 181 CGCGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAAATTAATTCCCAACCGCGTGGCAC 240
Db 2743 CGCGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAAATTAATTCCCAACCGCGTGGCAC 2802
Qy 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 2803 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 2862
Qy 301 ACGCGCGTCCCAATTTGTCGCGCGGATTAAATCTCGCGCGGATCACTCACTGGGTGCACGCG 360
Db 2863 ACGCGCGTCCCAATTTGTCGCGCGGATTAAATCTCGCGCGGATCACTCACTGGGTGCACGCG 2922
Qy 361 TGGTGGTTCGATGGTAGAACGAAAGCGCGCTCGAAGCTGTAAAGCGCGGTGCAAAATC 420
Db 2923 TGGTGGTTCGATGGTAGAACGAAAGCGCGCTCGAAGCTGTAAAGCGCGGTGCAAAATC 2982
Qy 421 TTCTCGCGCAACGCGTCAGTGGGCTGATTAATTAATCTATCCGCTGGATACCCAGGATGCCA 480
Db 2983 TTCTCGCGCAACGCGTCAGTGGGCTGATTAATTAATCTATCCGCTGGATACCCAGGATGCCA 3042
Qy 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db 3043 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 3102
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGATCGGCGGTGGAGCATC 600
Db 3103 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGATCGGCGGTGGAGCATC 3162
Qy 601 TGGTGGCATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 3163 TGGTGGCATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 3222
Qy 661 CGCGTCTCGCTCTGGCTGGGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 3223 CGCGTCTCGCTCTGGCTGGGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3282
Qy 721 CGAAACGGGAAAGCGGACGTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 3283 CGAAACGGGAAAGCGGACGTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 3342
Qy 781 ATGAGGGCATGTTCCGATCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGGCA 840
Db 3343 ATGAGGGCATGTTCCGATCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGGCA 3402
Qy 841 TGGCGCCATTACCGAGTCCGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 3403 TGGCGCCATTACCGAGTCCGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 3462
Qy 901 ACATACCGAAAGACAGCTCATGTTATATATCCGCGGTTAAACCAATCAAAACAGGATTTTC 960
Db 3463 ACATACCGAAAGACAGCTCATGTTATATATCCGCGGTTAAACCAATCAAAACAGGATTTTC 3522
Qy 961 GCCTCTGGGGCAAAACAGCGTGGAGCCGCTTGTGCAATCTCTCAGGCGCCAGGCGGTGA 1020
Db 3523 GCCTCTGGGGCAAAACAGCGTGGAGCCGCTTGTGCAATCTCTCAGGCGCCAGGCGGTGA 3582
Qy 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 3583 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 3642
Qy 1081 CGCAACCGCCCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGGCAGCACAGGTTT 1140
Db 3643 CGCAACCGCCCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGGCAGCACAGGTTT 3702
Qy 1141 CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTACTCATTAG 1200

Db 3703 CCGACTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 3762
QY 1201 GCAC 1204
Db 3763 GCAC 3766

RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerra, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3832 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-148-675A-2

Query Match 36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGGCAAAACCTTCCGGGTATGGCATGATAGGCCCGGAAGAGA 60
Db 3 CCGACACCATCGAATGGCGCAAAACCTTTCGGGTATGGCATGATAGGCCCGGAAGAGA 62

QY 61 GTCAATTGAGGGTGGTGAATGTGAACCCAGTAACTATACGATGTGCGAGATATGCCG 120
Db 63 GTCAATTGAGGGTGGTGAATGTGAACCCAGTAACTATACGATGTGCGAGATATGCCG 122

QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGTAACTATACGATGTGCGAGATATGCCG 180
Db 123 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGTAACTATACGATGTGCGAGATATGCCG 182

QY 181 CGCGGGAAGAAAGTGGAGCGCGATGGCGAGCTGAATTACATCCCAACGGCGTGGCAC 240
Db 183 CGCGGGAAGAAAGTGGAGCGCGATGGCGAGCTGAATTACATCCCAACGGCGTGGCAC 242

QY 241 AACAACTGGGGGCAACAGTCGTGTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300

Db 243 AACAACTGGGGGCAACAGTCGTGTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 302
QY 301 ACAGCGCGTCCAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTCGGTCGCGAGCG 360
Db 303 ACAGCGCGTCCAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTCGGTCGCGAGCG 362

QY 361 TGGTGGTGTGATGTTAGAACGAGCGCGCTGAGAGCTGTAAAGCGCGGTGCACAATC 420
Db 363 TGGTGGTGTGATGTTAGAACGAGCGCGCTGAGAGCTGTAAAGCGCGGTGCACAATC 422

QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCACTTAATCTCGCTGAGTACGAGATGCCA 480
Db 423 TTCTCGCGCAACCGCTCAGTGGGCTGATCACTTAATCTCGCTGAGTACGAGATGCCA 482

QY 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCTTAATTTCTTGATGTCTCTGACCAGA 540
Db 483 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCTTAATTTCTTGATGTCTCTGACCAGA 542

QY 541 CACCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 543 CACCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 602

QY 601 TGGTCCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCCAATTAAGTTCTCTCGG 660
Db 603 TGGTCCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCCAATTAAGTTCTCTCGG 662

QY 661 CGGCTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db 663 CGGCTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 722

QY 721 CGGAAACGGGAAGCGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
Db 723 CGGAAACGGGAAGCGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 782

QY 781 ATGAGGCAATGTTCCCATGCGATGCTGTTGCCAAACGATCAGATGCGCTGGGCGCAA 840
Db 783 ATGAGGCAATGTTCCCATGCGATGCTGTTGCCAAACGATCAGATGCGCTGGGCGCAA 842

QY 841 TGGCGCGCATTCAGGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 843 TGGCGCGCATTCAGGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 902

QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAACAGGATTTTC 960
Db 903 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAACAGGATTTTC 962

QY 961 GCCTGCTGGGGCAACACAGGTGGACCGTTGTCGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 963 GCCTGCTGGGGCAACACAGGTGGACCGTTGTCGCAACTCTCTCAGGCGCAGGCGGTGA 1022

QY 1021 AGGGCAATCAGCTGTTGCCCGCTCACTGTTGAAGAAAGAAACCAACCTTGGGCGCCCAATA 1080
Db 1023 AGGGCAATCAGCTGTTGCCCGCTCACTGTTGAAGAAAGAAACCAACCTTGGGCGCCCAATA 1082

QY 1081 CGAAACCCGCTCTCCCGCGGTTGGCGGATTCATTATGAGCTGGCAGCAGCAGGTTT 1140
Db 1083 CGAAACCCGCTCTCCCGCGGTTGGCGGATTCATTATGAGCTGGCAGCAGGTTT 1142

QY 1141 CCGGCTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTGAGTTAGTCTCATTTAG 1200
Db 1143 CCGGCTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTGAGTTAGTCTCATTTAG 1202

QY 1201 GCAC 1204
Db 1203 GCAC 1206

RESULT 9
US-08-487-283A-18/c
; Sequence 18, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.

APPLICANT: Matis, Louis A.
APPLICANT: Mueller, Bileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Russell P.
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Squinto, Stephen C.
APPLICANT: Thomas, James A.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
TITLE OF INVENTION: OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Cetris 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)776-1790
TELEFAX: (203)772-3655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: PET Trc SOS/NI
DESCRIPTION: prokaryotic expression vector
US-08-487-283A-18

Query Match 36.3%; Score 1199.2; DB 3; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCRAAACCTTTCCGGGTATGGCATGATAGCGCCGGGAAGA 60
DB 4810 CGGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGGCATGATAGCGCCGGGAAGA 4751
QY 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAACCTTATACGATCTCGCAGAGTATGCCG 120
DB 4750 GTCATTCAGGGTGGTGAATGTGAACACAGTAACCTTATACGATCTCGCAGAGTATGCCG 4691
QY 121 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCAGCCAGTTTCTCGAAAA 180
DB 4690 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCAGTTTCTCGAAAA 4631
QY 181 CGCGGAAAGTGAAGCGGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
DB 4630 CGCGGAAAGTGAAGCGGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 4571
QY 241 AACAACTCGCGGGCAACAGTGGTGTGATGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
DB 4570 AACAACTCGCGGGCAACAGTGGTGTGATGGGGTTGCCACCTCCAGTCTGGGCCCTGC 4511
QY 301 ACGGCGCGTCGCAAAATTTGTCGCGCGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGC 360

DB 4510 ACGCCCGTCGCAAAATTTGTCGCGCGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGC 4451
QY 361 TGGTGGTGTGATGTAGAACCAAGCGCGCTGGAAGCCTGTAAAGCGCGCGTGCACAATC 420
DB 4450 TGGTGGTGTGATGTAGAACCAAGCGCGCTGGAAGCCTGTAAAGCGCGCGTGCACAATC 4391
QY 421 TTTCGCGCAACCGCTCAGTGGGCTGATCAATACTATCCGCTGGATGACGAGATGCCA 480
DB 4390 TTTCGCGCAACCGCTCAGTGGGCTGATCAATACTATCCGCTGGATGACGAGATGCCA 4331
QY 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCGGGGTATTTCTTGTATGTCTCTGACCA 540
DB 4330 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCGGGGTATTTCTTGTATGTCTCTGACCA 4271
QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
DB 4270 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 4211
QY 601 TGGTCGCATTTGGGTCAACAGCAAAATCGGCTGTGTAGCGGGCCCATTTAAGTTCTGTCTCG 660
DB 4210 TGGTCGCATTTGGGTCAACAGCAAAATCGGCTGTGTAGCGGGCCCATTTAAGTTCTGTCTCG 4151
QY 661 CGGCTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
DB 4150 CGGCTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 4091
QY 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
DB 4090 CGGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTGCGATGTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
DB 4030 ATGAGGGCATCGTTCCCACTGCGATGTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 3971
QY 841 TGGCGGCAATTCACGAGTCCGGCTGCGGCTGGTGGCGGATATCTCGGTAGTGGGATAGC 900
DB 3970 TGGCGGCAATTCACGAGTCCGGCTGCGGCTGGTGGCGGATATCTCGGTAGTGGGATAGC 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCATCAAAAGGATTTTC 960
DB 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCATCAAAAGGATTTTC 3851
QY 961 GCCTGTGGGGCAACACGAGCTGCAACCGCTTGTGCTGCAACTCTCTCAGGGCCAGGCGTGA 1020
DB 3850 GCCTGTGGGGCAACACGAGCTGCAACCGCTTGTGCTGCAACTCTCTCAGGGCCAGGCGTGA 3791
QY 1021 AGGCAATCAGCTGTGGCGCTCTCACTGTGAAAGAAAGAAACCAACCTTGGGCGCCCAATA 1080
DB 3790 AGGCAATCAGCTGTGGCGCTCTCACTGTGAAAGAAAGAAACCAACCTTGGGCGCCCAATA 3731
QY 1081 CGCAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGAGTGGCGACGACAGGTTT 1140
DB 3730 CGCAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGAGTGGCGACGACAGGTTT 3671
QY 1141 CCCGACTGGAAGGGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 3670 CCCGACTGGAAGGGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 3611
QY 1201 GCAC 1204
DB 3610 GCAC 3607
RESULT 10
PCT-US96-05611A-21/c
; Sequence 21, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Lenardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Bileen Elliott

; APPLICANT: Nye, Steven H.
 ; APPLICANT: Pelfrey, Clara M.
 ; APPLICANT: Squinto, Stephen P.
 ; APPLICANT: Wilkins, James A.
 ; TITLE OF INVENTION: Modified Myelin Protein Molecules
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Maurice M. Klee
 ; STREET: 1951 Burr Street
 ; CITY: Fairfield
 ; STATE: Connecticut
 ; COUNTRY: USA
 ; ZIP: 06430
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
 ; COMPUTER: Macintosh Centris 610
 ; OPERATING SYSTEM: System 7
 ; SOFTWARE: Microsoft Word 6.0.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/05611A
 ; FILING DATE: 02-MAY-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/431,644
 ; FILING DATE: May 2, 1995
 ; APPLICATION NUMBER: 08/431,648
 ; FILING DATE: May 2, 1995
 ; APPLICATION NUMBER: 08/482,114
 ; FILING DATE: June 7, 1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Klee, Maurice M.
 ; REGISTRATION NUMBER: 30,399
 ; REFERENCE/DOCKET NUMBER: ALX-129
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (203) 255 1400
 ; TELEFAX: (203) 254 1101
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5248 base pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Double
 ; TOPOLOGY: Circular
 ; MOLECULE TYPE: Other nucleic acid
 ; DESCRIPTION: DET Trc S05/NI
 ; DESCRIPTION: Prokaryotic expression vector
 PCT-US96-05611A-21

Query Match 36.3%; Score 1199.2; DB 5; Length 5248;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CGGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
 DB 4810 CGGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 4751
 QY 61 GTCGAATTCAGGCGTGTGAATCTGAACACAGTAAAGTTATACGATGTCGAGAGTATCCG 120
 DB 4750 GTCGAATTCAGGCGTGTGAATCTGAACACAGTAAAGTTATACGATGTCGAGAGTATCCG 4691
 QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCGCAGCAGCTTCTGCGAAAA 180
 DB 4690 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCGCAGCAGCTTCTGCGAAAA 4631
 QY 181 CGCGGGAAGAGTGAAGCGCGATGGCGAGCTGAATTATCCCAACCGCGTGGCAC 240
 DB 4630 CGCGGGAAGAGTGAAGCGCGATGGCGAGCTGAATTATCCCAACCGCGTGGCAC 4571
 QY 241 AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGCCACTCCAGTCTGCCCTGC 300
 DB 4570 AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGCCACTCCAGTCTGCCCTGC 4511
 QY 301 ACGCGCGTGCAGAAATGTGCGCGGATTAATCTCGCGCGATCAACTCGGTGCCAGCG 360

Db 4510 ACGCGCGTGCAGAAATGTGCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGC 4451
 QY 361 TGGTGGTGTGATGATGAACGAAAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACATC 420
 Db 4450 TGGTGGTGTGATGATGAACGAAAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACATC 4391
 QY 421 TTCTCGCGCAACGCGTCAGTGGGTGATCAATTAATATCCGCTGGATGACACAGGATGCCA 480
 Db 4390 TTCTCGCGCAACGCGTCAGTGGGTGATCAATTAATATCCGCTGGATGACACAGGATGCCA 4331
 QY 481 TTGCTGTGGAAGCTGCCTGCACATTAATCTCGCGGTATTCTTGATGTCTCTGACAGA 540
 Db 4330 TTGCTGTGGAAGCTGCCTGCACATTAATCTCGCGGTATTCTTGATGTCTCTGACAGA 4271
 QY 541 CACCATCAACAGTATTATTCTCCATGAAGACGCTAGCGACTGGCGGTGGAGCATC 600
 Db 4270 CACCATCAACAGTATTATTCTCCATGAAGACGCTAGCGACTGGCGGTGGAGCATC 4211
 QY 601 TGGTCGCATTGGGTACACGCAATCGCGTGTAGCGGCGCCATTAAGTTCTGCTCGG 660
 Db 4210 TGGTCGCATTGGGTACACGCAATCGCGTGTAGCGGCGCCATTAAGTTCTGCTCGG 4151
 QY 661 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 Db 4150 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
 QY 721 CGGAAACGGGAAGGCGACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
 Db 4090 CGGAAACGGGAAGGCGACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAATGCTGA 4031
 QY 781 ATGAGGCGCATGTTCCCACTGCGATGCTGGTGGCCACAGATCAGATGGCGCTGGGCGCAA 840
 Db 4030 ATGAGGCGCATGTTCCCACTGCGATGCTGGTGGCCACAGATCAGATGGCGCTGGGCGCAA 3971
 QY 841 TGGCGCGCATTAACCGAGTCCGGCTCGCGGTGGTGGGATATCTCGTAGTGGGATACG 900
 Db 3970 TGGCGCGCATTAACCGAGTCCGGCTCGCGGTGGTGGGATATCTCGTAGTGGGATACG 3911
 QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTAAACACCATCAAAACAGGATTTTC 960
 Db 3910 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTAAACACCATCAAAACAGGATTTTC 3851
 QY 961 GCCTGCTGGGGCAACACAGCGTGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
 Db 3850 GCCTGCTGGGGCAACACAGCGTGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA 3791
 QY 1021 AGGCGAATCAGCTGTGGCGGTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
 Db 3790 AGGCGAATCAGCTGTGGCGGTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 3731
 QY 1081 CGCAAAACCCCTCTCCCGCGGTGGCGGATTCATTAATGAGTGGCGACGAGTTT 1140
 Db 3730 CGCAAAACCCCTCTCCCGCGGTGGCGGATTCATTAATGAGTGGCGACGAGTTT 3671
 QY 1141 CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGCTCACTCATTTAG 1200
 Db 3670 CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGCTCACTCATTTAG 3611
 QY 1201 GCAC 1204
 Db 3610 GCAC 3607

RESULT 11
 US-10-263-103-35/c
 ; Sequence 35, Application US/10263103
 ; Patent No. 6821723
 ; GENERAL INFORMATION:
 ; APPLICANT: AVENTIS PASTEUR
 ; APPLICANT: Chevallier, Michel
 ; APPLICANT: El Habib, Raphaelle
 ; APPLICANT: Krell, Tino
 ; APPLICANT: Sodoyer, Regis

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pET-cer
US-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGC	AAACCTTTCCGGGTATG	GCATGATAGCGCCGGAAGAGA	60
DB	4622	CGGACACCATCGAATGGCG	CAAAACCTTTCCGGGTATG	GCATGATAGCGCCGGAAGAGA	4563
QY	61	GTCAATTACAGGTGGTGA	ATGTGAACACGATTAAC	CGTATACGATGTCGACAGTATGCGG	120
DB	4562	GTCAATTACAGGTGGTGA	ATGTGAACACGATTAAC	CGTATACGATGTCGACAGTATGCGG	4503
QY	121	GTGTCCTTATCAGACCG	CTTTCCCGCGTGTGAAC	CAGCGCCACCGACGTTTCTGCGAAAA	180
DB	4502	GTGTCCTTATCAGACCG	CTTTCCCGCGTGTGAAC	CAGCGCCACCGACGTTTCTGCGAAAA	4443
QY	181	CGCGGAAAAAGTGAAG	CGGCGATGGCGGAGCTG	AAATTACATTTCCCAACCGCGTGGCAC	240
DB	4442	CGCGGAAAAAGTGAAG	CGGCGATGGCGGAGCTG	AAATTACATTTCCCAACCGCGTGGCAC	4383
QY	241	AACAACTGCGGGCAAA	CAGTCTGTTGCTGATT	GGCGTTGGCCACCTCCAGTCTGGCCCTGC	300
DB	4382	AACAACTGCGGGCAAA	CAGTCTGTTGCTGATT	GGCGTTGGCCACCTCCAGTCTGGCCCTGC	4323
QY	301	ACGCGCGCTCGCAAA	TTGTCGCGCGGATTAAT	CTCGCGCGGATCAACTGGGTGGCAGCG	360
DB	4322	ACGCGCGCTCGCAAA	TTGTCGCGCGGATTAAT	CTCGCGCGGATCAACTGGGTGGCAGCG	4263
QY	361	TGGTGGTTCGATGTG	AGAACGAGCGGCTGAA	AGCGCTGTAAGCGGCGTGGCAATC	420
DB	4262	TGGTGGTTCGATGTG	AGAACGAGCGGCTGAA	AGCGCTGTAAGCGGCGTGGCAATC	4203
QY	421	TTCTCGCGCAACGCTG	AGTGGGCTGATCATT	TAATCTCCGCTGGATGACCAAGGATGCCA	480
DB	4202	TTCTCGCGCAACGCTG	AGTGGGCTGATCATT	TAATCTCCGCTGGATGACCAAGGATGCCA	4143
QY	481	TTGCTGTGGAAGCTG	CGCTGCACATAATGTT	CCGGCGTTATTTCTTGATGTTCTTGACCCAGA	540
DB	4142	TTGCTGTGGAAGCTG	CGCTGCACATAATGTT	CCGGCGTTATTTCTTGATGTTCTTGACCCAGA	4083
QY	541	CACCCATCAACAGTAT	TATTTCTCCATGAAG	ACGGTACCGGCTGGGCGTGGAGCATC	600
DB	4082	CACCCATCAACAGTAT	TATTTCTCCATGAAG	ACGGTACCGGCTGGGCGTGGAGCATC	4023
QY	601	TGGTGCATTTGGTC	CACAGCAAAATCGC	GTGTTAGCGGGCCCATTAAGTTCGTCTCGG	660
DB	4022	TGGTGCATTTGGTC	CACAGCAAAATCGC	GTGTTAGCGGGCCCATTAAGTTCGTCTCGG	3963
QY	661	CGCGTCTGGCTGGC	TGGCTGCAATAATCT	CTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	3962	CGCGTCTGGCTGGC	TGGCTGCAATAATCT	CTCACTCGCAATCAAAATTCAGCCGATAG	3903
QY	721	CGGAACGGGAACGG	CATGTGACATGTCG	GTGTTTCAACAAACCATGCAAAATGCTGA	780
DB	3902	CGGAACGGGAACGG	CATGTGACATGTCG	GTGTTTCAACAAACCATGCAAAATGCTGA	3843
QY	781	ATGAGGGCATCGT	CCCCCTGCGATGCT	GTGTTGCCAACGATCAGATGGCGTGGCGGCA	840
DB	3842	ATGAGGGCATCGT	CCCCCTGCGATGCT	GTGTTGCCAACGATCAGATGGCGTGGCGGCA	3783

QY	841	TGCGCGCATTTACCG	AGTCCGGGCTGCGG	CTGGTGGGATATCTCGGTAGTGGATACG	900
DB	3782	TGCGCGCATTTACCG	AGTCCGGGCTGCGG	CTGGTGGGATATCTCGGTAGTGGATACG	3723
QY	901	ACGATACCGAAGAC	GAGCTCATGTTATAT	CCCGCGTTAACCCACCATCAACAGGATTTTC	960
DB	3722	ACGATACCGAAGAC	GAGCTCATGTTATAT	CCCGCGTTAACCCACCATCAACAGGATTTTC	3663
QY	961	GCCTGCTGGGCAAA	CCAGCGCTGGAC	CCCTTGTGCAACTCTCTCAGGCGCCGCGTGA	1020
DB	3662	GCCTGCTGGGCAAA	CCAGCGCTGGAC	CCCTTGTGCAACTCTCTCAGGCGCCGCGTGA	3603
QY	1021	AGGCAATCAGCTGT	TGGCCCTCTCACTG	GTGAAAGAAAAACCACTTGGCGCCCAATA	1080
DB	3602	AGGCAATCAGCTGT	TGGCCCTCTCACTG	GTGAAAGAAAAACCACTTGGCGCCCAATA	3543
QY	1081	CGCAAAACCGCTCT	CCCCCGCGTTGG	CGGATTCATTAATGCAGCTGGCAGCAGGTTT	1140
DB	3542	CGCAAAACCGCTCT	CCCCCGCGTTGG	CGGATTCATTAATGCAGCTGGCAGCAGGTTT	3483
QY	1141	CCCAGCTGAAAG	CGGCGCAGTGAG	CGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
DB	3482	CCCAGCTGAAAG	CGGCGCAGTGAG	CGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG	3423
QY	1201	GCAC	1204		
DB	3422	GCAC	3419		

RESULT 12

US-08-929-967-1/c
; Sequence 1, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Ruppert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 4810 CGGACACCATCGAATGGCGCAAAACCTTTCCGGTATGGCATATAGCGCCCGGAAGAGA 4751
QY 61 GTCATTTAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTGCGAGTATGCCG 120
Db 4750 GTCATTTAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTGCGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCGCACCGCTTTCTGCGAAAA 180
Db 4690 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCGCACCGCTTTCTGCGAAAA 4631
QY 181 CGCGGGAAGAGTGAAGCGCGCATGGCGAGCTGAATACATTCCTCAACCGCGTGGCAC 240
Db 4630 CGCGGGAAGAGTGAAGCGCGCATGGCGAGCTGAATACATTCCTCAACCGCGTGGCAC 4571
QY 241 AACAACTGGCGGGCAACAGCTGCTGTGATGCGCGTGGCGACCTCCAGTCTGGCCCTGC 300
Db 4570 AACAACTGGCGGGCAACAGCTGCTGTGATGCGCGTGGCGACCTCCAGTCTGGCCCTGC 4511
QY 301 ACGGCGCGTGCAGAAATGTGCGCGCGATTAAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 4510 ACGGCGCGTGCAGAAATGTGCGCGCGATTAAATCTCGCGCGCATCAACTGGGTGCCAGCG 4451
QY 361 TGGTGTGTGCGATGGTAGAACGAGCGGCTGAAGCTTGAAGCGCGGTGCAATC 420
Db 4450 TGGTGTGTGCGATGGTAGAACGAGCGGCTGAAGCTTGAAGCGCGGTGCAATC 4391
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCGCATCAACTGGGTGCCAGCG 480
Db 4390 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCGCATCAACTGGGTGCCAGCG 4331
QY 481 TTGCTGTGAGCTGCGCTGCATTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
Db 4330 TTGCTGTGAGCTGCGCTGCATTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 4271
QY 541 CACCCATCAACAGTATTTTCTCCATGAAGAGCGGTACGCGCTGGCGGTGGAGATC 600
Db 4270 CACCCATCAACAGTATTTTCTCCATGAAGAGCGGTACGCGCTGGCGGTGGAGATC 4211
QY 601 TGGTGCATTTGGGTGCACGAAATGCGCGCTTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 4210 TGGTGCATTTGGGTGCACGAAATGCGCGCTTTAGCGGGCCCAATTAAGTTCTGTCTCGG 4151
QY 661 CGCGTCTGGTCTGGCTGCGCATAAATCTCACTCGCAATCAATTCAGCGGATAG 720
Db 4150 CGCGTCTGGTCTGGCTGCGCATAAATCTCACTCGCAATCAATTCAGCGGATAG 4091
QY 721 CGGAAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAATGCTGA 780
Db 4090 CGGAAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAATGCTGA 4031
QY 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTTGCCAAGCATCAGATGGCGGTGGCGCAA 840
Db 4030 ATGAGGGCATGTTCCCACTGGCATGCTGGTTGCCAAGCATCAGATGGCGGTGGCGCAA 3971
QY 841 TGCGCGCATTAACGAGTCCGGGCTGCGCGTGTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 3970 TGCGCGCATTAACGAGTCCGGGCTGCGCGTGTGGTGGGATATCTCGGTAGTGGGATAG 3911
QY 901 ACGATACCGAAGACAGCTATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 3851
QY 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db 3850 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 3791
QY 1021 AGGGCAATCAGCTGTGGCGGTCTCACTGGTGAAGAAAAAACACCCCTGGCGCCCAATA 1080
Db 3790 AGGGCAATCAGCTGTGGCGGTCTCACTGGTGAAGAAAAAACACCCCTGGCGCCCAATA 3731
QY 1081 CGCAACCGCGCTCTCCCGCGCGTGGCCGATTCATTAATGCACTGGCACGACGAGTTT 1140

Db 3730 CCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATGCACTGGCACGACAGTTT 3671
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Db 3670 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607
RESULT 13
US-09-702-705-785/c
; Sequence 785, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Badgur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-785

Query Match 36.1%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 4618
QY 61 GTCATTTAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTGCGAGTATGCCG 120
Db 4617 GTCATTTAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTGCGAGTATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCGCACCGTTTCTGCGAAAA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCGCACCGTTTCTGCGAAAA 4498
QY 181 CGCGGGAAGAGTGAAGCGCGCATGGCGAGCTGAATTCATTCCTCAACCGCGTGGCAC 240
Db 4497 CGCGGGAAGAGTGAAGCGCGCATGGCGAGCTGAATTCATTCCTCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAACAGCTGCTGTGATGCGCGTGGCGACCTCCAGTCTGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAACAGCTGCTGTGATGCGCGTGGCGACCTCCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGTGCAGAAATGTGCGCGCGATTAAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 4377 ACGCGCGTGCAGAAATGTGCGCGCGATTAAATCTCGCGCGCATCAACTGGGTGCCAGCG 4318
QY 361 TGGTGTGTGCGATGGTAGAACGAGCGCGTGAAGCTTGAAGCGCGGTGCACAATC 420
Db 4317 TGGTGTGTGCGATGGTAGAACGAGCGCGTGAAGCTTGAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATGCACTGGCACGACGAGTTT 480

Db 4257 TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 4198
Qy 481 TTGCTGTGGAAGCTGCCCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 540
Db 4197 TTGCTGTGGAAGCTGCCCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 4138
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 4078
Qy 601 TGGTGCATTTGGTTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCTGTTCTCGG 660
Db 4077 TGGTGCATTTGGTTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCTGTTCTCGG 4018
Qy 661 CGCGTCTGCGTCTGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 3898
Qy 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCAA 840
Db 3897 ATGAGGGCATCTTCCCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCAA 3838
Qy 841 TCGCGCCATTACCGAGTCCGGCTCGCGTTCGGTTCGGGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCATTACCGAGTCCGGCTCGCGTTCGGTTCGGGATATCTCGGTAGTGGGATACG 3778
Qy 901 ACGATACCGAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAACAGGATTTTC 960
Db 3777 ACGATACCGAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAACAGGATTTTC 3718
Qy 961 GCCTGCTGGGCAACACCGCTGACCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGCAACACCGCTGACCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 3658
Qy 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3657 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 3598
Qy 1081 CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTATGCACTGGCACACAGGTTT 1140
Db 3597 CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTATGCACTGGCACACAGGTTT 3538
Qy 1141 CCGGACTGGAAGCGGCGAGTGCAGCGCAACGCAATTAATGTCAGTTAGTCACTCATTAG 1200
Db 3537 CCGGACTGGAAGCGGCGAGTGCAGCGCAACGCAATTAATGTCAGTTAGTCACTCATTAG 3478
Qy 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 14

US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGCCCGCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGGCGCAAAACCTTTTCGCGGTATGGCATGATAGCCCGCGGAAGAGA 4618
Qy 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGATGTCGAGAGTATCGG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGATGTCGAGAGTATCGG 4558
Qy 121 GTGTCCTTTATCAGACCCGTTTCCCGCGTGTGAACCCAGGCCAGCCAGCTTTCTGCGAAAA 180
Db 4557 GTGTCCTTTATCAGACCCGTTTCCCGCGTGTGAACCCAGGCCAGCCAGCTTTCTGCGAAAA 4498
Qy 181 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 4438
Qy 241 AACAACTGGCGGGCAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGGCCCTGC 4378
Qy 301 ACGGCGCTGCGCAAAATTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 4377 ACGGCGCTGCGCAAAATTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 4318
Qy 361 TGGTGGTTCGATGATGAAGCGCGCTGCAAGCTGTAAGCGCGGTGCAACAATC 420
Db 4317 TGGTGGTTCGATGATGAAGCGCGCTGCAAGCTGTAAGCGCGGTGCAACAATC 4258
Qy 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGGATGCCA 480
Db 4257 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGGATGCCA 4198
Qy 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 540
Db 4197 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 4138
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 4078
Qy 601 TGGTGCATTTGGTTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCTGTTCTCGG 660
Db 4077 TGGTGCATTTGGTTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCTGTTCTCGG 4018
Qy 661 CGCGTCTGCGTCTGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 3898
Qy 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCAA 840
Db 3897 ATGAGGGCATCTTCCCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCAA 3838
Qy 841 TCGCGCCATTACCGAGTCCGGCTCGCGTTCGGTTCGGGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCATTACCGAGTCCGGCTCGCGTTCGGTTCGGGATATCTCGGTAGTGGGATACG 3778

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell updates/sec

Title: US-09-765-555B-18
Perfect score: 3300
Sequence: 1 ccgacacatcgatggtgc.....acgacgtccggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3300	100.0	3300	6	AX202430	Sequence
2	3232.8	98.0	3300	6	AX202429	Sequence
3	3226.4	97.8	3300	6	AX202426	Sequence
4	3226.4	97.8	3300	6	AX202427	Sequence
5	3216.8	97.5	3300	6	AX202428	Sequence
6	2705.8	82.0	3300	6	AX284144	Sequence
7	2705.8	82.0	8101	6	AX172306	Sequence
8	2701	81.8	6648	6	AX377531	Sequence
9	2700	81.8	7475	6	A91965	Sequence
10	2700	81.8	7475	6	AR031992	Sequence
11	2700	81.8	7475	6	AR207294	Sequence
12	2698	81.8	9191	6	AX377532	Sequence
13	2676.2	81.1	6724	12	AF097412	Sequence
14	2636.8	79.9	6806	6	AX378208	Sequence
15	2631	79.7	6706	12	AF031088	Sequence
16	2598.2	78.7	6748	12	AF031813	Sequence
17	1609.4	48.8	5558	6	CQ794769	Sequence
18	1474	44.7	4700	6	CQ846980	Sequence
19	1462.4	44.3	4921	6	AX768201	Sequence

20	1462.4	44.3	4935	6	AX768200	Sequence
21	1462.4	44.3	4945	6	AX768202	Sequence
22	1462.4	44.3	4951	6	AX768203	Sequence
23	1432.4	43.4	6755	12	AF140577	Integrati
24	1417.6	43.0	1511	6	AX431576	Sequence
25	1417.6	43.0	4297	12	AY219685	Expressio
26	1417.6	43.0	4359	12	AY219686	Expressio
27	1417.6	43.0	4419	12	AY219688	Expressio
28	1417.6	43.0	4855	12	AY219689	Expressio
29	1417.6	43.0	4968	12	AY219687	Expressio
30	1417.6	43.0	6956	12	AY219682	Shuttle e
31	1417.6	43.0	7018	12	AY219683	Shuttle e
32	1417.6	43.0	7509	12	AY219684	Shuttle e
33	1400.8	42.4	4566	12	AF177933	Cloning v
34	1400.8	42.4	4635	12	AF177932	Cloning v
35	1400.8	42.4	6399	12	AY302759	Expressio
36	1400.8	42.4	6587	12	AY289542	Expressio
37	1362	41.3	5201	6	AR493834	Sequence
38	1362	41.3	5201	6	AR493835	Sequence
39	1329.2	40.3	4806	12	AF050464	Expressio
40	1238.4	37.5	4803	12	AY243506	Cloning v
41	1238.4	37.5	5024	6	CQ846977	Sequence
42	1237.2	37.5	1922	6	AX478528	Sequence
43	1204	36.5	4506	12	AB116367	Expressio
44	1204	36.5	4557	6	E13740	DNA sequenc
45	1204	36.5	4557	6	AR373251	Sequence

ALIGNMENTS

RESULT 1
AX202430
LOCUS AX202430 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 18 from Patent WO0152620.
ACCESSION AX202430
VERSION AX202430.1 GI:15392178

KEYWORDS
synthetic construct
other sequences; artificial sequences.
ORGANISM
synthetic construct
other sequences; artificial sequences.

REFERENCE
1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 18 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)

FEATURES
source
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-Ap3 and zinc finger protein ZFPap3"

ORIGIN	Query Match	100.0%;	Score 3300;	DB 6;	Length 3300;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3300;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	CGACACCATCGAATGGTGC	AAACCTTCGCGGTATGCGCATGATGAGCCCGGAGAGA	60	
Db	1	CGACACCATCGAATGGTGC	AAACCTTCGCGGTATGCGCATGATGAGCCCGGAGAGA	60	
QY	61	GTCATTTCAGGTGGTGAATG	GAACACAGTAACGTTATACCATGTGCGCAGAGTATGCCG	120	
Db	61	GTCATTTCAGGTGGTGAATG	GAACACAGTAACGTTATACCATGTGCGCAGAGTATGCCG	120	
QY	121	GTGTCCTTATCAGACCGTTT	CCCCCGTGGTGAACACGAGCCAGCCACGCTTTCTGCGAAAA	180	
Db	121	GTGTCCTTATCAGACCGTTT	CCCCCGTGGTGAACACGAGCCAGCCACGCTTTCTGCGAAAA	180	
QY	181	CGCGGAAAAAGTGGAAAGCG	CGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGGTGGCAC	240	

Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAAATTAATTCCTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAAAACAGTCGTGTCTGATTTGGCGTTGGCCACCTCCAGTCTTGCCCTTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCGTGTCTGATTTGGCGTTGGCCACCTCCAGTCTTGCCCTTGC 300
Qy 301 ACGCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Qy 361 TGTGTGTTCATCGGTAGAACGAAGCGCGTCCGAGCTGTAAAGCGCGGTGCACATC 420
Db 361 TGTGTGTTCATCGGTAGAACGAAGCGCGTCCGAGCTGTAAAGCGCGGTGCACATC 420
Qy 421 TTCTCGGCAACGGGTCAAGTGGGTGATCAATTAATCTCCCTGGATGACACGAGTGCCA 480
Db 421 TTCTCGGCAACGGGTCAAGTGGGTGATCAATTAATCTCCCTGGATGACACGAGTGCCA 480
Qy 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGGACTGGCGTGGAGCATC 600
Qy 601 TGTGTGATTTGGGTCAACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGTGTGATTTGGGTCAACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCGATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCGATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGAGCTGGAGTCCCATGTCGGGTTTCAACAAACCATGCAATGTCTGA 780
Db 721 CGGAACGGGAAGCGAGCTGGAGTCCCATGTCGGGTTTCAACAAACCATGCAATGTCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTCGCATGTGTTGCCAACGATCAGATGGCGTGGCGGCA 840
Db 781 ATGAGGGCATCGTTCCCACTCGCATGTGTTGCCAACGATCAGATGGCGTGGCGGCA 840
Qy 841 TGC CGCGCATTAACGAGTCCGGGCTGCGGTGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGC CGCGCATTAACGAGTCCGGGCTGCGGTGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAACAGAGATTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAACAGAGATTTTC 960
Qy 961 GCCTGTGGGGCAAAACGAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACGAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGCAATACAGCTGTGGCGGTCTCACTGTGTAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGCAATACAGCTGTGGCGGTCTCACTGTGTAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Qy 1081 CGCAACCGGCTCTCCCGCGGTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAACCGGCTCTCCCGCGGTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 1200
Qy 1201 GCACAAATCTCATGTTTGAACGCTTATCATGCACTGCAAGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAACGCTTATCATGCACTGCAAGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGTATGGCTGTGAGGTGTGTAATCACTGCATAATTCG 1320

Db 1261 TCAGGAGCCATCGGAAGCTGTGTATGCTGTGTCAGGTGTAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGGGGCACTCCCGTTCTCGATAATGTTTTTTTGGCCGACATCATACCGTT 1380
Db 1321 TGTGCTCAAGGGGCACTCCCGTTCTCGATAATGTTTTTTTGGCCGACATCATACCGTT 1380
Qy 1381 CTGCAAAATTTCTGAAATGAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Db 1381 CTGCAAAATTTCTGAAATGAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Qy 1441 ATTGTGAGCGGATACAAATTTCAACAGGAAAACAGCCAGTCCGTTTAGGTGTTTTTCAGGA 1500
Db 1441 ATTGTGAGCGGATACAAATTTCAACAGGAAAACAGCCAGTCCGTTTAGGTGTTTTTCAGGA 1500
Qy 1501 GCATCTCACCAACAAAGGACCATGATTTATGAAACTGAAGAAGTAACTGGTAAATCTGG 1560
Db 1501 GCATCTCACCAACAAAGGACCATGATTTATGAAACTGAAGAAGTAACTGGTAAATCTGG 1560
Qy 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCCCGACAAAGCTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCCCGACAAAGCTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTACACCGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTACACCGCAAGCTGATGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
Qy 1921 GAGATCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGGTACTTCACTGCGCGCTGATGCTGCTGACCGGGGTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGGTACTTCACTGCGCGCTGATGCTGCTGACCGGGGTATGCTTCAAG 2040
Qy 2041 TATGAAAAACGGCAAGTACGACATTAAGAGCTGGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGAGCTGGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGATGCTGATTTAAAAAACAACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGATGCTGATTTAAAAAACAACATGAATGAGACACCGATTAC 2160
Qy 2161 TCCATCGCAAGCTGCTTTTAAAGGGAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAAGCTGCTTTTAAAGGGAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGTCCAACTGACACACGACAAAGTGAATTTATGGTGTAAACGGGTACTGCCGACCTTC 2280
Db 2221 GCATGTCCAACTGACACACGACAAAGTGAATTTATGGTGTAAACGGGTACTGCCGACCTTC 2280
Qy 2281 AAGGTCACCATCCAAAACCGTTCTGGTGGCGTCTGAGCGCAGGTATTACGGCCCGAGT 2340
Db 2281 AAGGTCACCATCCAAAACCGTTCTGGTGGCGTCTGAGCGCAGGTATTACGGCCCGAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTCTG 2400

Qy	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAGATCCACGCTATTGGCGCCACCATGGAAACGCCACAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAGATCCACGCTATTGGCGCCACCATGGAAACGCCACAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGCTTTCTTGATGTCGCTGCGTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGCTTTCTTGATGTCGCTGCGTACTGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Qy	2641	AACAAACAACAATAAACAATAACAACACCTCGGGATCGAGGGAAGGATTTTCAGAAATC	2700
Db	2641	AACAAACAACAATAAACAATAACAACACCTCGGGATCGAGGGAAGGATTTTCAGAAATC	2700
Qy	2701	GGATCCTCTTCTGTCGCGCCAGCGGCCCTTCGAGCCCGGGGAGAGCCCTATGCTTGT	2760
Db	2701	GGATCCTCTTCTGTCGCGCCAGCGGCCCTTCGAGCCCGGGGAGAGCCCTATGCTTGT	2760
Qy	2761	CCGGAATGTGTAAGTCTTTCAGCCAGAGCAGTCCCTGGTGGCCACACAGCGTACCCAC	2820
Db	2761	CCGGAATGTGTAAGTCTTTCAGCCAGAGCAGTCCCTGGTGGCCACACAGCGTACCCAC	2820
Qy	2821	ACGGGTGAAGAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAGAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGTCCAGCAAC	2880
Qy	2881	CTGGTGGCCCATCAACGCACTCATACTGCGGAGAGCCATACAAATGTCAGAAATGTGGC	2940
Db	2881	CTGGTGGCCCATCAACGCACTCATACTGCGGAGAGCCATACAAATGTCAGAAATGTGGC	2940
Qy	2941	AAGTCTTTTCAGCAGTCCAGCAACTGTGTGGCCACCAACGTACTCACACGGGGAGAAG	3000
Db	2941	AAGTCTTTTCAGCAGTCCAGCAACTGTGTGGCCACCAACGTACTCACACGGGGAGAAG	3000
Qy	3001	CCCTATGCTTGCCTGGATGTGTAAGTCTTTCAGCAGCAGTGGCTCTTGGTTAGACAC	3060
Db	3001	CCCTATGCTTGCCTGGATGTGTAAGTCTTTCAGCAGCAGTGGCTCTTGGTTAGACAC	3060
Qy	3061	CAGCGTACCCACACGGGTGAAACCGGTATAATGCCAGAGTGGCGCAATCTTTTAGC	3120
Db	3061	CAGCGTACCCACACGGGTGAAACCGGTATAATGCCAGAGTGGCGCAATCTTTTAGC	3120
Qy	3121	CAGCGCGCCACCTGGAAAGCCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT	3180
Db	3121	CAGCGCGCCACCTGGAAAGCCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT	3180
Qy	3181	CCAGAAATGGCAAGTCTTCTCAACTTCAGCACTTGGTCCGTACCAACGTAATC	3240
Db	3181	CCAGAAATGGCAAGTCTTCTCAACTTCAGCACTTGGTCCGTACCAACGTAATC	3240
Qy	3241	ACCGGTAAAGAACTAGTGGCCAGCGCCAGTACCGGTACGAGTTCGCGACTACGCT	3300
Db	3241	ACCGGTAAAGAACTAGTGGCCAGCGCCAGTACCGGTACGAGTTCGCGACTACGCT	3300
RESULT 2	AX202429	Sequence 17 from Patent WO0152620.	
LOCUS	AX202429	3300 bp	DNA linear PAT 30-AUG-2001
DEFINITION	AX202429		
ACCESSION	AX202429		
VERSION	AX202429.1	GI:15392177	
KEYWORDS		synthetic construct	
SOURCE		synthetic construct	
ORGANISM		other sequences; artificial sequences.	
REFERENCE	1		

AUTHORS	Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.									
TITLE	Methods and compositions to modulate expression in plants									
JOURNAL	Patent: WO 0152620-A 17 26-JUL-2001; The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL DISCOVERY, INC. (CA)									
FEATURES	Location/Qualifiers									
source	1..3300									
	/organism="synthetic construct"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:32630"									
	/notes="Partial sequence of pMal-m4 and zinc finger protein ZFPm4"									
ORIGIN										
Query Match	98.0%; Score 3232.8; DB 6; Length 3300;									
Best Local Similarity	98.7%; Pred. No. 0;									
Matches 3258; Conservative	0; Mismatches 42; Indels 0; Gaps 0;									
Qy	1	CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGCGATCATAGCGCCCGGAAGAGA	60							
Db	1	CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGCGATCATAGCGCCCGGAAGAGA	60							
Qy	61	GTCAATTCCAGGTGGTGAATGTGAACCCAGTAAACGTTATACGATGTCGACAGATGATGCCG	120							
Db	61	GTCAATTCCAGGTGGTGAATGTGAACCCAGTAAACGTTATACGATGTCGACAGATGATGCCG	120							
Qy	121	GTGTCCTTTATCAGACCGTTTCCGCGGTGTAACACGAGCCAGCCACGTTTCTGCGAAAAA	180							
Db	121	GTGTCCTTTATCAGACCGTTTCCGCGGTGTAACACGAGCCAGCCACGTTTCTGCGAAAAA	180							
Qy	181	CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGGTGGCAC	240							
Db	181	CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGGTGGCAC	240							
Qy	241	AACAACTGGCGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTCGCCCTGC	300							
Db	241	AACAACTGGCGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTCGCCCTGC	300							
Qy	301	ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG	360							
Db	301	ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG	360							
Qy	361	TGTTGGTTCGATGGTAGAAAGAGCGGCTCGAAGCTGTGTAAGCGGCGGTGCACAATC	420							
Db	361	TGTTGGTTCGATGGTAGAAAGAGCGGCTCGAAGCTGTGTAAGCGGCGGTGCACAATC	420							
Qy	421	TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACACGAGTGC	480							
Db	421	TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACACGAGTGC	480							
Qy	481	TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA	540							
Db	481	TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA	540							
Qy	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600							
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600							
Qy	601	TGTTCCGATTTGGGTTCACCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTC	660							
Db	601	TGTTCCGATTTGGGTTCACCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTC	660							
Qy	661	CGGCTCTGCGCTGCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720							
Db	661	CGGCTCTGCGCTGCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720							
Qy	721	CGGAACGGGAAGCGACTGGAGTGCATGATCTCCGTTTTTCAACAAACCATGCAAAATGCTGA	780							
Db	721	CGGAACGGGAAGCGACTGGAGTGCATGATCTCCGTTTTTCAACAAACCATGCAAAATGCTGA	780							
Qy	781	ATGAGGCGATCGTTCCCACTCGGATGCTGTTGGCAACGATCAGATGGCGCTGGGCGCAA	840							
Db	781	ATGAGGCGATCGTTCCCACTCGGATGCTGTTGGCAACGATCAGATGGCGCTGGGCGCAA	840							

QY	841	TGCGGCCATTACCGAGTCCGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGGCCATTACCGAGTCCGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCACTCAAAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCACTCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAACACAGCGTGGACCGCTGCTGCAACTCTCTCAGGCCACGCGGTGA	1020
Db	961	GCCTGCTGGGCAACACAGCGTGGACCGCTGCTGCAACTCTCTCAGGCCACGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTGGCCGATTCATTAAATGCGAGTGGCAACAGGTTT	1140
Db	1081	CGCAACCGCTCTCCCGCGGTGGCCGATTCATTAAATGCGAGTGGCAACAGGTTT	1140
QY	1141	CCGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAACGGTGCAACCAATGTTCTGGCG	1260
Db	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAACGGTGCAACCAATGTTCTGGCG	1260
QY	1261	TCAGGACGCAATCGGAGCTGTGATGCTGTGCGAGTGTGATTAATCACTGCATAATTCG	1320
Db	1261	TCAGGACGCAATCGGAGCTGTGATGCTGTGCGAGTGTGATTAATCACTGCATAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACTCCCGTCTCTGGATAATGTTTTTCGCGCGACATCAACCGTT	1380
Db	1321	TGTCGCTCAAGCGCACTCCCGTCTCTGGATAATGTTTTTCGCGCGACATCAACCGTT	1380
QY	1381	CTGGCAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA	1440
Db	1381	CTGGCAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTACACAGGAACAGCCAGTCCGTTTAGTGTGTTTACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTACACAGGAACAGCCAGTCCGTTTAGTGTGTTTACGA	1500
QY	1501	GCACTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGTAACTGGTAATCTGG	1560
Db	1501	GCACTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGTAACTGGTAATCTGG	1560
QY	1561	ATTAAACCGGATAAAGGCTATAACGGTCTCGCTGAAGTAAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACCGGATAAAGGCTATAACGGTCTCGCTGAAGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTACCGTTGAGCATCCGATATACTGGAAGAGAAATTCACACAGTT	1680
Db	1621	ACCGGAATTAAGTACCGTTGAGCATCCGATATACTGGAAGAGAAATTCACACAGTT	1680
QY	1681	CGGCAACTGCGCATGCGCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGGCAACTGCGCATGCGCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCGCTGTTGGCTGAATCAACCGGACAAAGCGTTTCCAGGACAAAGCTGTAT	1800
Db	1741	GCTCAATCTGCGCTGTTGGCTGAATCAACCGGACAAAGCGTTTCCAGGACAAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGATGCGTACGTTAACTCAACGGCAAGCTGATTCTTACCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGATGCGTACGTTAACTCAACGGCAAGCTGATTCTTACCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGTCGCGAACCCGCAAAAACCTGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGTCGCGAACCCGCAAAAACCTGGAA	1920

QY	1921	GAGATCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAAACCGTACTTTCACCTGCGCCCTGATTCCTGCTGACGGGGTTATGCGGTTCAAG	2040
Db	1981	CTGCAAGAAACCGTACTTTCACCTGCGCCCTGATTCCTGCTGACGGGGTTATGCGGTTCAAG	2040
QY	2041	TATGAAAAACGGCAAGTACGACATTTAAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAAAACGGCAAGTACGACATTTAAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG	2100
QY	2101	GCTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACACATGATGACAGACCGGATTAC	2160
Db	2101	GCTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACACATGATGACAGACCGGATTAC	2160
QY	2161	TCCATCGCAGAAGCTGCTCTTAAATAAGGCGAAACAGCGATGACCATCAACGGCGCGTGG	2220
Db	2161	TCCATCGCAGAAGCTGCTCTTAAATAAGGCGAAACAGCGATGACCATCAACGGCGCGTGG	2220
QY	2221	GCATGCTCAACATCGACACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCGACCTTC	2280
Db	2221	GCATGCTCAACATCGACACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCGACCTTC	2280
QY	2281	AAGGTCACCATCTCAAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACGGCGCCAGT	2340
Db	2281	AAGGTCACCATCTCAAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACGGCGCCAGT	2340
QY	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTCG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTCG	2400
QY	2401	GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	2460
QY	2461	TTGGCGAAAGATCCAGTATTGCGCGCACCATCGAAAAACCCAGAAAGGTGAATTCATG	2520
Db	2461	TTGGCGAAAGATCCAGTATTGCGCGCACCATCGAAAAACCCAGAAAGGTGAATTCATG	2520
QY	2521	CCGAAACATCCGCGAGATGTCGGTTCCTGATGTCGGTGGTGGTGGTGGTGGTGGTGGTGG	2580
Db	2521	CCGAAACATCCGCGAGATGTCGGTTCCTGATGTCGGTGGTGGTGGTGGTGGTGGTGGTGG	2580
QY	2581	GCAGCGGCTGTCAGACTGTCGATGAAGCCCTGAAGAGCGCGAGACTAATTCGAGCTCG	2640
Db	2581	GCAGCGGCTGTCAGACTGTCGATGAAGCCCTGAAGAGCGCGAGACTAATTCGAGCTCG	2640
QY	2641	AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAAAGGATTTTCAGAAATTC	2700
Db	2641	AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAAAGGATTTTCAGAAATTC	2700
QY	2701	GGATCTCTCTCTGTCGCGCGCGCGCTGAGCGCGGGGAGAGCCCTATGCTTGT	2760
Db	2701	GGATCTCTCTCTGTCGCGCGCGCGCTGAGCGCGGGGAGAGCCCTATGCTTGT	2760
QY	2761	CCGGAATGTTGTAAGTCTTTCAGCCAGACAGCTCCCTGCTGGCCACCGAGCTATCCAC	2820
Db	2761	CCGGAATGTTGTAAGTCTTTCAGCCAGACAGCTCCCTGCTGGCCACCGAGCTATCCAC	2820
QY	2821	ACGGGTGAAAAACCGGTATTAATGCGGCAATCTTTTAGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAAACCGGTATTAATGCGGCAATCTTTTAGCCAGTCCAGCAAC	2880
QY	2881	CTGGTGGCCCATCAACGCACTCATCTGCGGAGAGCCCATCAAAATGTCAGAAATGTCG	2940
Db	2881	CTGGTGGCCCATCAACGCACTCATCTGCGGAGAGCCCATCAAAATGTCAGAAATGTCG	2940
QY	2941	AAGTCTTTTTCAGCGAGTCCAGCAACCTGGTGGCGCCCAACAGTACTCACACCGGGGAGAG	3000
Db	2941	AAGTCTTTTTCAGTGTATGTCGTGATCTTGGAGGACCAACAGTACTCACACCGGGGAGAG	3000
QY	3001	CCCTATGCTGTGTCGGGAATGTTGTAAGTCTTTCAGCACCAAGTGGCTCTTGGTTAGACAC	3060

|||||
3001 CCCTATGTTGTCGGGATGTGGTAAGTCTCTCTCTCAGAGCTCTCACCTGGTGGCCAC 3060
Qy
3061 CAGCGTACCCACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTACG 3120
Db
3061 CAGCGTACCCACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTACG 3120
Qy
3121 CAGCGCGCCACCTGGAGCGGCATCAACGCACCTATCTACTGGGAGAACCATACAAATGT 3180
Db
3121 CGCAGCGTAACCTGGTGGCGCATCAACGCACCTATCTACTGGGAGAACCATACAAATGT 3180
Qy
3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTCAGGCAACTTGGTCCGTACCAAGTACTCAC 3240
Db
3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTCAGGCAACTTGGTCCGTACCAAGTACTCAC 3240
Qy
3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCGCAGTACCGGTACGAGCTTCCGAGCTACGCT 3300
Db
3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCGCAGTACCGGTACGAGCTTCCGAGCTACGCT 3300

RESULT 3
AX202426
LOCUS AX202426 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 14 from Patent WO0152620.
ACCESSION AX202426
VERSION AX202426.1 GI:15392173
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.

REFERENCE
1 Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
METHODS AND COMPOSITIONS TO MODULATE EXPRESSION IN PLANTS
PATENT: WO 0152620-A 14 26-JUL-2001;
THE SCRIPPS RESEARCH INSTITUTE (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)

FEATURES
source
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m1 and zinc finger protein
ZFpml"

ORIGIN
Query Match 97.8%; Score 3226.4; DB 6; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCATAAACTTTTCGGCGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCATAAACTTTTCGGCGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCATTCAGGCTGTGATGTGAACAGTAACCTTATACGATTCGCGAGTATGCG 120
Db 61 GTCATTCAGGCTGTGATGTGAACAGTAACCTTATACGATTCGCGAGTATGCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCCAGCCAGCTTTCGCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCCAGCCAGCTTTCGCGAAAA 180
Qy 181 CCGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAAATTTACATTCCTCCACCGCGTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAAATTTACATTCCTCCACCGCGTGGCAC 240
Qy 241 ACAAATCGCGGGCAACAGTCTGCTGATTTGGGTGGCGACCTCCAGTCTGGCCCTGC 300
Db 241 ACAAATCGCGGGCAACAGTCTGCTGATTTGGGTGGCGACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCCCGTTCGCAAAATGTTCGGCGGATTAATCTCGGCCGATCAATCTGGGTGCCAGCG 360
Db 301 ACGCCCGTTCGCAAAATGTTCGGCGGATTAATCTCGGCCGATCAATCTGGGTGCCAGCG 360

Qy 361 TGTGTTGTCGATGGTAGAACGAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACATC 420
Db 361 TGTGTTGTCGATGGTAGAACGAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACATC 420
Qy 421 TTTCTCGCGCAACCGGTTCAGTGGGCTGATCATTAACTATCCCGTGGATCACCAGATGCCA 480
Db 421 TTTCTCGCGCAACCGGTTCAGTGGGCTGATCATTAACTATCCCGTGGATCACCAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAAATGTTCCGGGTTATTTCTTGATGTCCTGACCAAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAAATGTTCCGGGTTATTTCTTGATGTCCTGACCAAGA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACCGACTGGCGGTGGAGCATC 600
Qy 601 TGTGTCGATTCGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGTGTCGATTCGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Qy 661 CGGCTCTGCGTCTGGCTGGGCATAAATATCTCACTCGCAATCAAATTCAGCGGATAG 720
Db 661 CGGCTCTGCGTCTGGCTGGGCATAAATATCTCACTCGCAATCAAATTCAGCGGATAG 720
Qy 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAAGCATCAGATGGCGTGGCGGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAAGCATCAGATGGCGTGGCGGCAA 840
Qy 841 TGGCGCCATTACCGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAGC 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAGC 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC 960
Qy 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGTCCTCTCATGTTGTAAGAAAAAACCACTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTCCTCTCATGTTGTAAGAAAAAACCACTTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGCGAGCTGCGACAGGTTT 1140
Db 1081 CGCAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGCGAGCTGCGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGAGTTAGTCTACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGAGTTAGTCTACTCATTTAG 1200
Qy 1201 GCACAAATTCATGTTTGCAGAGCTTATCATCGACTGCAAGTGGTGAACCAATGCTTGGCG 1260
Db 1201 GCACAAATTCATGTTTGCAGAGCTTATCATCGACTGCAAGTGGTGAACCAATGCTTGGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGATTAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCTGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Qy 1441 ATTGTGAGGGGATTAACAATTTTCAACAGGAAACGCCAGTCCGTTAGTGTGTTTCAAGA 1500

Db	1441	ATTGTGAGCGATAA	CAATTTCA	CACAGGAAAC	CGCCAGTCC	GGTTTAGGTGTTT	CA	CGA	1500	
Qy	1501	GCATTTACCA	CAAGGACCA	TAGATTAT	GAAAACT	GAAGAAGTA	AACTGGT	TAATCTGG	1560	
Db	1501	GCATTTACCA	CAAGGACCA	TAGATTAT	GAAAACT	GAAGAAGTA	AACTGGT	TAATCTGG	1560	
Qy	1561	ATTAACGGCGATA	AAAGGCTAA	ACGGTCTCG	CTCCTGAA	CTCGTAAGAA	TTTCGAGAA	AGAT	1620	
Db	1561	ATTAACGGCGATA	AAAGGCTAA	ACGGTCTCG	CTCCTGAA	CTCGTAAGAA	TTTCGAGAA	AGAT	1620	
Qy	1621	ACCGGAATTA	AAAGTCAC	CGTTTGAC	ATCCGGAT	ATAA	CTGGAAGAGAA	ATTTCCCA	1680	
Db	1621	ACCGGAATTA	AAAGTCAC	CGTTTGAC	ATCCGGAT	ATAA	CTGGAAGAGAA	ATTTCCCA	1680	
Qy	1681	CGGCAACTGG	CGATGCGCTG	ACATTTAT	CTTCTGGG	GCACAC	CGCTTTGGTGG	GTAC	1740	
Db	1681	CGGCAACTGG	CGATGCGCTG	ACATTTAT	CTTCTGGG	GCACAC	CGCTTTGGTGG	GTAC	1740	
Qy	1741	GCTCAATCTG	GCCTGTTG	GGCTGAA	ATCACCCCG	GACAAAGCG	TTCCAGGACA	AGCTGTAT	1800	
Db	1741	GCTCAATCTG	GCCTGTTG	GGCTGAA	ATCACCCCG	GACAAAGCG	TTCCAGGACA	AGCTGTAT	1800	
Qy	1801	CCGTTTACTG	GGATGCGGT	ACGTTA	CAACGG	CAAGCTGAT	GTGCTTACC	CGATCGCTGTT	1860	
Db	1801	CCGTTTACTG	GGATGCGGT	ACGTTA	CAACGG	CAAGCTGAT	GTGCTTACC	CGATCGCTGTT	1860	
Qy	1861	GAAGCGTTAT	CTGCTGAT	TTATAACAA	AGATCTGCTG	CCGACACCG	CCAAACCTGG	AA	1920	
Db	1861	GAAGCGTTAT	CTGCTGAT	TTATAACAA	AGATCTGCTG	CCGACACCG	CCAAACCTGG	AA	1920	
Qy	1921	GAGATCCG	CGCTGGATAA	AGAACTGAA	AGCGAA	AGGTAA	AGCGCGCTGAT	GTTC	1980	
Db	1921	GAGATCCG	CGCTGGATAA	AGAACTGAA	AGCGAA	AGGTAA	AGCGCGCTGAT	GTTC	1980	
Qy	1981	CTGCAAGAAC	CGGTACTT	CACTGG	CGCTGAT	TGCTGCTG	ACGGGGTTAT	GCTTCAAG	2040	
Db	1981	CTGCAAGAAC	CGGTACTT	CACTGG	CGCTGAT	TGCTGCTG	ACGGGGTTAT	GCTTCAAG	2040	
Qy	2041	TATGAAAC	CGCAAGTAC	GCAT	TAAAGAC	GTGGCGTGAT	AACGCTG	CGCGGAAAGCG	2100	
Db	2041	TATGAAAC	CGCAAGTAC	GCAT	TAAAGAC	GTGGCGTGAT	AACGCTG	CGCGGAAAGCG	2100	
Qy	2101	GGTCTGAC	CTTCTGGTTG	ACCTGAT	TAAAAA	CAACACAT	GAATGAC	ACACCGGATTAC	2160	
Db	2101	GGTCTGAC	CTTCTGGTTG	ACCTGAT	TAAAAA	CAACACAT	GAATGAC	ACACCGGATTAC	2160	
Qy	2161	TCCATCGCA	AGCTGCTTTT	ATAAAGG	CGAAACAG	CGATGAC	CAACCGCGCGTGG	2220		
Db	2161	TCCATCGCA	AGCTGCTTTT	ATAAAGG	CGAAACAG	CGATGAC	CAACCGCGCGTGG	2220		
Qy	2221	GCATGCTC	CAACATCG	ACACAG	CAAGTGA	ATTATGGTGT	ATAACGGTACTG	CGACCTTC	2280	
Db	2221	GCATGCTC	CAACATCG	ACACAG	CAAGTGA	ATTATGGTGT	ATAACGGTACTG	CGACCTTC	2280	
Qy	2281	AAAGGTCA	ACCATCCAA	ACCGTTT	CTGGCGTCTG	AGCGCAGGTAT	TAAACGCGCGCAGT	2340		
Db	2281	AAAGGTCA	ACCATCCAA	ACCGTTT	CTGGCGTCTG	AGCGCAGGTAT	TAAACGCGCGCAGT	2340		
Qy	2341	CCGAA	CAAGAGCTG	GC	CAAAAGAG	TTTCTCG	AAAACTATCTG	CTGACTGATGAAGTCTG	2400	
Db	2341	CCGAA	CAAGAGCTG	GC	CAAAAGAG	TTTCTCG	AAAACTATCTG	CTGACTGATGAAGTCTG	2400	
Qy	2401	GAAGCGGTT	TAATAAG	ACAAACCG	CTGGTGGT	CGTAGCGCT	GAACTTTC	ACGAGGAAGAG	2460	
Db	2401	GAAGCGGTT	TAATAAG	ACAAACCG	CTGGTGGT	CGTAGCGCT	GAACTTTC	ACGAGGAAGAG	2460	
Qy	2461	TTGGCGAA	AGATCC	ACGTATTG	CGCGCAC	CAATCG	GAAGAA	CGCCAGAAAGTGAATCATG	2520	
Db	2461	TTGGCGAA	AGATCC	ACGTATTG	CGCGCAC	CAATCG	GAAGAA	CGCCAGAAAGTGAATCATG	2520	
Qy	2521	CCGAA	CATCCG	CAGATG	CTCGCTTT	CTGGTATG	CGCTGCTG	CTGCTGCTGATTC	2580	
Db	2521	CCGAA	CATCCG	CAGATG	CTCGCTTT	CTGGTATG	CGCTGCTG	CTGCTGCTGATTC	2580	

ORIGIN		ZFPm2"		Db				
Query Match 97.8%; Score 3226.4; DB 6; Length 3300; Best Local Similarity 98.6%; Pred. No. 0; Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;								
Qy	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60	Qy	961	GCCTGCTGGGGCAAAAC	CGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGCGTGA	1020
Db	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60	Db	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAAGAAAAACCACTCTGGCGCCCAATA	1080	
Qy	61	GTCAATTCCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGATATGCCG	120	Qy	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAAGAAAAACCACTCTGGCGCCCAATA	1080	
Db	61	GTCAATTCCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGATATGCCG	120	Db	1081	CGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140	
Qy	121	GTGTCCTTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAAA	180	Qy	1081	CGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140	
Db	121	GTGTCCTTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAAA	180	Db	1081	CGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140	
Qy	181	CGCGGAAAAAGTGAACGGCGGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC	240	Qy	1141	CCGACTTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200	
Db	181	CGCGGAAAAAGTGAACGGCGGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC	240	Db	1141	CCGACTTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200	
Qy	241	AACAACTGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGCGCCTGC	300	Qy	1201	GCAAAATTTCTGTTTGGACAGCTTATCATCGACTGACGCGTGACCAATGTTCTTGCGG	1260	
Db	241	AACAACTGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGCGCCTGC	300	Db	1201	GCAAAATTTCTGTTTGGACAGCTTATCATCGACTGACGCGTGACCAATGTTCTTGCGG	1260	
Qy	301	ACGCGCCGTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360	Qy	1261	TCAGGACGCAATCGGAAAGCTGTGTATGGCTGTGCAGGTCTGTAATCACTGCATAATTCG	1320	
Db	301	ACGCGCCGTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360	Db	1261	TCAGGACGCAATCGGAAAGCTGTGTATGGCTGTGCAGGTCTGTAATCACTGCATAATTCG	1320	
Qy	361	TGTTGGTGTGATGTGAAGCAAGCGGGTGAAGCGCTGTAAGCGCGGTGCAAAATC	420	Qy	1321	TGTCGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATTAACGGTT	1380	
Db	361	TGTTGGTGTGATGTGAAGCAAGCGGGTGAAGCGCTGTAAGCGCGGTGCAAAATC	420	Db	1321	TGTCGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATTAACGGTT	1380	
Qy	421	TTCTCGCGCAACGGTCAAGTGGGCTGATTAATCTCCGCTGATCCGCTGGATGACAGGATGCCA	480	Qy	1381	CTGCAAAATTTCTGAAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTTTCG	1440	
Db	421	TTCTCGCGCAACGGTCAAGTGGGCTGATTAATCTCCGCTGATCCGCTGGATGACAGGATGCCA	480	Db	1381	CTGCAAAATTTCTGAAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTTTCG	1440	
Qy	481	TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540	Qy	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA	1500	
Db	481	TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540	Db	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA	1500	
Qy	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600	Qy	1501	GCACCTTCCAAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTCG	1560	
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600	Db	1501	GCACCTTCCAAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTCG	1560	
Qy	601	TGTTGCGCATTTGGGTACACGAGAAATCGCGCTGTGAGCGGCGCCATTAAGTTCTGTCGCG	660	Qy	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620	
Db	601	TGTTGCGCATTTGGGTACACGAGAAATCGCGCTGTGAGCGGCGCCATTAAGTTCTGTCGCG	660	Db	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620	
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	Qy	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT	1680	
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	Db	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT	1680	
Qy	721	CGGAAACGGGAAGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA	780	Qy	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCAACGACCGCTTTGGTGGCTAC	1740	
Db	721	CGGAAACGGGAAGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA	780	Db	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCAACGACCGCTTTGGTGGCTAC	1740	
Qy	781	ATGAGGGCATCGTTCCCACTGGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA	840	Qy	1741	GCTCAATCTGGCCTGTTGGCTGNAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT	1800	
Db	781	ATGAGGGCATCGTTCCCACTGGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA	840	Db	1741	GCTCAATCTGGCCTGTTGGCTGNAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT	1800	
Qy	841	TGCGGCGCATTAACGAGTCCGGGCTGGGTGGTGGGATATCTCGGTAGTGGGATACG	900	Qy	1801	CGGTTTACCTGGGATGCCGTACGTTAACAACGCAAGCTGATTTGCTTACCCGATCGCTGTT	1860	
Db	841	TGCGGCGCATTAACGAGTCCGGGCTGGGTGGTGGGATATCTCGGTAGTGGGATACG	900	Db	1801	CGGTTTACCTGGGATGCCGTACGTTAACAACGCAAGCTGATTTGCTTACCCGATCGCTGTT	1860	
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAACAGGATTTTC	960	Qy	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA	1920	
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAACAGGATTTTC	960	Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA	1920	
Qy	961	GCCTGCTGGGGCAAAACCAAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGCGCGTGA	1020	Qy	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980	
Db				Db	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980	


```
QY 2101 GGTCTGAACCTTCTCGTTGAACTGATTAATAAACAACACATGAATGACAGACCGAATTAC 2160
Db 2101 GGTCTGAACCTTCTCGTTGAACTGATTAATAAACAACACATGAATGACAGACCGAATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTAAATAAAGCGGAAACAGAGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTAAATAAAGCGGAAACAGAGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACACAGCAAAAGTGAATTATGGTGAACCGTACTCCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACAGCAAAAGTGAATTATGGTGAACCGTACTCCGACCTTC 2280
QY 2281 AAGGTTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAAGTATTTGGCGCCACCATGGAACACGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAAGTATTTGGCGCCACCATGGAACACGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGGCTTCTGGTATGCGGTGATCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGCTTCTGGTATGCGGTGATCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGAGACTGTGATGAAGCCCTGAAAGACGGCAGACTAATTCAGTCTG 2640
Db 2581 GCCAGCGTCTGAGACTGTGATGAAGCCCTGAAAGACGGCAGACTAATTCAGTCTG 2640
QY 2641 AACAAACAACAATACATTAACAACAACCTCGGGATGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATACATTAACAACAACCTCGGGATGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCACGCGCCCTCCAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCACGCGCCCTCCAGCCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGAAATGTGTAAAGTCTTCTCAGCAGCAGAGCTCCCTGTGTCGCCACACGCGTACCCAC 2820
Db 2761 CCGAAATGTGTAAAGTCTTCTCAGCAGCTCTCACCTGTGTGCGCCACACGCGTACCCAC 2820
QY 2821 ACGGTGAACCAACCGTATAATGCCAGAGTGGCGGAAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAACCAACCGTATAATGCCAGAGTGGCGGAAATCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCATCAACGCACTCATCTGGGAGAGGCATACAATGTCCAGAAATGGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATCTGGGAGAGGCATACAATGTCCAGAAATGGGC 2940
QY 2941 AAGTCTTTAGCCAGTCCAGCAACCTGTGGCCACCAAGCTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCCGGCACCACAGTACTCACACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTGTAAAGTCTTTCAGACACAGTGGTCTCTTGGTTAGACAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGTAAAGTCTTTCAGCCGACGCGATACCTGGTGGCGCAC 3060
QY 3061 CAGCGTACCACACAGCGGTGAACCAACCGTATAAATGCGCAGAGTGGCGCAATCTTTTACG 3120
Db 3061 CAGCGTACCACACAGCGGTGAACCAACCGTATAAATGCGCAGAGTGGCGCAATCTTTTACG 3120
QY 3121 CAGCGGCCCACTGTGAACCGCATCAACGCACTCATCTGCGGAGAGCGCATCAAAATGT 3180
Db 3121 CAGCGGCCCACTGTGGCCAGCCATCAACGCACTCATCTGCGGAGAGCGCATCAAAATGT 3180
```

```
QY 3181 CCAGAATGTGGCAAGTCTTTTCTCAACTTTCAGGCAACTTTGGTCCGTACCAACGTAACAC 3240
Db 3181 CCAGAATGTGGCAAGTCTTTTCTCGGTCGTGACAAATCTGTCGGCACCAACGTAACAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGGCGGCGGACGATACCCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGGCGGCGGACGATACCCGTACGACGTTCCGGACTACGCT 3300

RESULT 5
AX202428 LOCUS 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 16 from Patent WO0152620.
ACCESSION AX202428
VERSION AX202428.1 GI:15392176
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas, C. F., Stege, J. T., Guan, X. and Dalmia, B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 16 JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source 1..3300
location/qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Partial sequence of pMal-m3 and zinc finger protein ZFPm3"
```

ORIGIN

```
Query Match 97.5%; Score 3216.8; DB 6; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAGAGA 60
QY 61 GTCATTCAGGTTGGTGAATGTGAACACCAAGTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCATTCAGGTTGGTGAATGTGAACACCAAGTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCAGCTTCTTCCGAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCAGCTTCTTCCGAAA 180
QY 181 CCGGGGAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTCATATCCCAACCGCGTGGCAC 240
Db 181 CCGGGGAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTCATATCCCAACCGCGTGGCAC 240
QY 241 AACACTGGCGGCAACAGTCTGTTGATGGCGTTGCCACTCCAGTCTGGCCCTGC 300
Db 241 AACACTGGCGGCAACAGTCTGTTGATGGCGTTGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCCGCTGGCAAAATGTCGGCGCATTAATATCTCGCGCGATCACTGGGTGGCAGCG 360
Db 301 ACGCCGCTGGCAAAATGTCGGCGCATTAATATCTCGCGCGATCACTGGGTGGCAGCG 360
QY 361 TGGTGTGTGATGGTAGAACGAAGCGGCGTGAAGCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGGTGTGTGATGGTAGAACGAAGCGGCGTGAAGCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTATCCGTGGATGACAGAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTATCCGTGGATGACAGAGATGCCA 480
QY 481 TTGCTGTGAAGCTGCTGCACCTAAATGTTCCGCGGTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACCTAAATGTTCCGCGGTATTTCTTGATGTCTCTGACCAGA 540
```


QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGATCGGTCACCGACAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGATCGGTCACCGACAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACCTGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACCTGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCNA 840
QY 841 TGGCGCCATTACCGAGTCCGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCCATTACCGAGTCCGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGCGCGCTCTCACTGTTGAAAGAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCGCTCTCACTGTTGAAAGAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACGACAGGTTT 1140
QY 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGAGTTAGTCACTCATTTAG 1200
Db 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGAACCAATGAGTTAGTCACTCATTTAG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGAACCAATGAGTTAGTCACTCATTTAG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAAATCACTGCAATTTAG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAAATCACTGCAATTTAG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAGGA 1500
Db 1441 ATTGTAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAGGA 1500
QY 1501 GCACCTTCAACAAAGGACCATAGATTTATGAAAACCTGAAGAGGTAATTCAGGATTC 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATTTATGAAAACCTGAAGAGGTAATTCAGGATTC 1560
QY 1561 ATTTAAGCGGATAAAGGCTTAAACCGTCTCGCTGAAAGTTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTTAAACCGTCTCGCTGAAAGTTCGAGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCAACCGTTGACATCCCGATAACTGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGACATCCCGATAACTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 GCGCAAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATCAAAAACGCAAGTACGACATTAAGAACGCTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATCAAAAACGCAAGTACGACATTAAGAACGCTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGGTACCTGATTAATAAAACAAACACATGATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGGTACCTGATTAATAAAACAAACACATGATGACAGACACCGATTAC 2160
QY 2161 TCCATCCGAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGCTCAACCATCAAAACCGTTCTGTTGGCTGTGAGCGCAGGTATTAACCGCCAGT 2340
Db 2281 AAGGCTCAACCATCAAAACCGTTCTGTTGGCTGTGAGCGCAGGTATTAACCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTGCGGCCCAACCTGGAAGAAACGCGCCAGAAAGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTGCGGCCCAACCTGGAAGAAACGCGCCAGAAAGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTCGCTGATGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTCGCTGATGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAAATAACAAACCAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACCAACAAACAAATAACAAACCAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
QY 2701 GGATCCTCTTCTGTTGGCCACAGCGGCCCTCTGAGCGCGGGGAGAGACCCCTATGCTGT 2760

```
Db 2701 GGAATCCTCTTCTGTGGCCAGCGCGCTCGAGCCGGGAGAACCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTCTAGCCAGAGCAGTCCCTGGTGGCCACACAGGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTCTAGCCAGTCTGGTGGCCACCTGGTGGCCACAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGCCAGCGGTCC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCTACTGGGAGAGCCCATCAAAATGTCAGAGTGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCTACTGGGAGAGCCCATCAAAATGTCAGAGTGTGGC 2940
Qy 2941 AAGTCTTTTTCAGCAGTCCAGCAACCTGGTGGCCACCAACGTAAGTCTTTCAGCAGTCCAGCAAC 3000
Db 2941 AAGTCTTTTTCAGCAGTCCAGCCTGGTGGCCACCAACGTAAGTCTTTCAGCAGTCCAGCAAC 3000
Qy 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCAGTCCAGTGGTCTTTCAGTGGTACAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCAGTCCAGTGGTCTTTCAGTGGTGGC 3060
Qy 3061 CAGCGTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGC 3120
Db 3061 CAGCGTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGT 3120
Qy 3121 CAGCGCGCCACCTTGGAAAGCGCATCAACGCACTCTACTGGCGAGAGCCATCAAAATGT 3180
Db 3121 GACTGCCGCGACCTTGTTCGCGCATCAACGCACTCTACTGGCGAGAGCCATCAAAATGT 3180
Qy 3181 CCAGATGTGGCAAGTCTTTCACACTTCAGGCAACTTGGTCCGTCACCAACGTAAGTAC 3240
Db 3181 CCAGATGTGGCAAGTCTTTCACCAATCCAGCCATCTCGTCCGCGCAACGTAAGTAC 3240
Qy 3241 ACCGTTAAAAAATAGTGGCGAGCGCGCCAGTACCGTACGAGCTTCGGACTACGCT 3300
Db 3241 ACCGTTAAAAAATAGTGGCGAGCGCGCCAGTACCGTACGAGCTTCGGACTACGCT 3300

RESULT 6
AX284144
LOCUS
DEFINITION Sequence 1 from Patent WO0178756.
ACCESSION AX284144
VERSION AX284144.1 GI:17044832
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS Wiedersanders,B. and Maubach,G.
TITLE Agent for postoperative use after the removal of bone tumours
JOURNAL Patent: WO 0178756-A 1 25-OCT-2001;
Depuy Biotech Jena GmbH (DE)
FEATURES
source
1. .7373
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Expressionsvektor pMal2cchbnp2"
1528. .3453
/notes="unnamed protein product; /produkt=Fusionsprotein
MBP Cystatin C BMP-2"
/codon_start=1
/transl_table=11
/protein_id="CAD12160.1"
/db_xref="GI:17044833"
/translators="MKIBEGKLVINGDKGYNGLAIEVKKFKDTGIKVTVEHPDKL
EEKFPQVATGDDPDIIFWADRFGGYAQSGLLAETPDRAFDQKLYPTFTWADRYNG
KLATYAPREALSLIYNKDIILPNPKTWEEIIPALDKELKAKGKSLMFLNQBYPFWP
LIAADGGYAFKPYNGKDIIDKGVNDAGAKGLTFLVDLTKNKHMDTDYSIAEAF
NKGETAWINGPNAWSNIDTSKYNVGVTVLPTFKGQSPKFPFVGLSAGINAASPNKEL
```

```
AKPLENLLTDRGLEAVNKDKPLGAVLAKSBEELAKDPRIATMENAKQKEMPMI
PQMSAFYAVRTAVINASGRQIVDEALDKAQINSSNNNNNNNNNNNNNNNNNN
SSSPGPRVLUGGPMDSVEEGRRALDFAVGEYNKASNDMHSRALQVVRARQIV
AGVNYFLDVBELGRTCTKTPQNLNDNCPFDHQLKXKAFCSQIYAVPQGTWTLKS
TCDDASASHHHHGHAEATAQAKHQRKLSKSKRHPLYYVDFSDVGNWDIVAPGVH
AFYCHBCPPPLADHLNSTWHAIVQTLVNSVNSKI PKACCVPTLSAISMLYDENEK
VLKNYQDMVVEGCGR"
misc_feature 2689
/note="Faktor Xa Protease /position=388 (AMINO
ACIDsequenz)"
3067. .3108
misc_feature
/note="/spacer zwischen Cystatin C und BMP-2"
ORIGIN
Query Match 82.0%; Score 2705.8; DB 6; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGTATGGCATATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGTATGGCATATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTCAGGTTGGTGAATGTCAAAACCACTAAGTTATACGATGTGCGAGAGTATCGG 120
Db 61 GTCAATTCAGGTTGGTGAATGTCAAAACCACTAAGTTATACGATGTGCGAGAGTATCGG 120
Qy 121 GTGCTCTTATCAGACCGTTTCCGCGGTGTGAACAGGCGCAGCGTTCCTGCGAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCGCGGTGTGAACAGGCGCAGCGTTCCTGCGAAAA 180
Qy 181 CCGGGAAAAAAGTGGAGCGCGGATGGCGGAGCTGAATTAATTCCCAACCGCGTGGCAC 240
Db 181 CCGGGAAAAAAGTGGAGCGCGGATGGCGGAGCTGAATTAATTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGGCAACAGTCTGTGATTTGGCGGTGCGACCTCAAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGGCAACAGTCTGTGATTTGGCGGTGCGACCTCAAGTCTGGCCCTGC 300
Qy 301 ACSCGCGCTCGCAAAATGTGCGGCGGATTAATACTCGCGCGGATCAACTGGGTGGCAGG 360
Db 301 ACSCGCGCTCGCAAAATGTGCGGCGGATTAATACTCGCGCGGATCAACTGGGTGGCAGG 360
Qy 361 TGTGTTGTTCGATGGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGTGTTGTTCGATGGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCGTCACTGAGTGGCTGATCAATTAATATCCGCTGGATGACAGAGTGC 480
Db 421 TTCTCGCGCAACCGCGTCACTGAGTGGCTGATCAATTAATATCCGCTGGATGACAGAGTGC 480
Qy 481 TTGCTGTGGAAGCTGCTGCACATAATTTTCGCGCGTTATTTCTTGTATGCTCTGACGAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACATAATTTTCGCGCGTTATTTCTTGTATGCTCTGACGAGA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGTGTCGATTTGGTTCACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660
Db 601 TGTGTCGATTTGGTTCACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Qy 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGTTTTCACAAACCAACCATGCAATCTGA 780
Db 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGTTTTCACAAACCAACCATGCAATCTGA 780
Qy 781 ATGAGGGCATCTGTTCCCACTCGGATGCTGTTGCCACCATCAGATGGGCGCTGGGGCGAA 840
Db 781 ATGAGGGCATCTGTTCCCACTCGGATGCTGTTGCCACCATCAGATGGGCGCTGGGGCGAA 840
```

QY 841 TCGCGCCATTACCGAGTCCGGCTTGGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
DB 841 TCGCGCCATTACCGAGTCCGGCTTGGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACACCATCAAAACAGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACACCATCAAAACAGATTTTC 960
QY 961 GCCTCTGGGGCAAAACACGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGGCGTGA 1020
DB 961 GCCTCTGGGGCAAAACACGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGGCGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGGAAACCAACCTTGGCGGCCAATA 1080
DB 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGGAAACCAACCTTGGCGGCCAATA 1080
QY 1081 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGAGTGGCAGCAGGTTT 1140
DB 1081 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGAGTGGCAGCAGGTTT 1140
QY 1141 CCCGACTCGAAAGCGGAGTGGCGCAACGCAATTAATGTAGTCTCACTCACTTAG 1200
DB 1141 CCCGACTCGAAAGCGGAGTGGCGCAACGCAATTAATGTAGTCTCACTCACTTAG 1200
QY 1201 GCACAATTTCTATGTTGACAGCTTATCATCGACTGCAAGGCAACCAATGCTTCTGGCG 1260
DB 1201 GCACAATTTCTATGTTGACAGCTTATCATCGACTGCAAGGCAACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGCATTCGAAAGCTGGTATGGCTGTCAGGTCGTAATCACTGCATATTCG 1320
DB 1261 TCAGGCGCATTCGAAAGCTGGTATGGCTGTCAGGTCGTAATCACTGCATATTCG 1320
QY 1321 TGTGCTCAAGCGCATCTCCGCTCTCGGATAATGTTTTTGGCGGACATCAACGGTT 1380
DB 1321 TGTGCTCAAGCGCATCTCCGCTCTCGGATAATGTTTTTGGCGGACATCAACGGTT 1380
QY 1381 CTGGCAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
DB 1381 CTGGCAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTAGGGGATACAAATTTACACAGGAAACAGCCAGTCCGTTAGTGTTCACGA 1500
DB 1441 ATTGTAGGGGATACAAATTTACACAGGAAACAGCCAGTCCGTTAGTGTTCACGA 1500
QY 1501 GCACTTTCACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAACTCGG 1560
DB 1501 GCACTTTCACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAACTCGG 1560
QY 1561 ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGAATAACTGGAAAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGAATAACTGGAAAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGCGCTTCTGGTGTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGCGCTTCTGGTGTAC 1740
QY 1741 GCTCAATCTGGGCTGTTGGCTGAATATCACCCGGCAAAACGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGGCTGTTGGCTGAATATCACCCGGCAAAACGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTAGCTTAAACCGCAAGCTGATTCGCTTACCCGATCGCTGT 1860
DB 1801 CCGTTTACCTGGATGCGGTAGCTTAAACCGCAAGCTGATTCGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCGCCAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCGCCAAACCTGGGAA 1920

QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAACCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATAAAGAACTGAACCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAAACGTAATCTTCCACCTGGCGCTGATTTGCTGTCGACGGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAAACGTAATCTTCCACCTGGCGCTGATTTGCTGTCGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAACCTGGCGGTGGATAACGTTGGCGGCAAGCG 2100
DB 2041 TATGAAAAACGCAAGTACGACATTAAGAACCTGGCGGTGGATAACGTTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGCTGACCTGATTAATAAACAACAACATCAATGACAGACCGATTAC 2160
DB 2101 GGTCTGACCTTCTGCTGCTGACCTGATTAATAAACAACAACATCAATGACAGACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGCAACAGCAATGACCATCAACGGCCCGTGG 2220
DB 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGCAACAGCAATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACACGCAAAAGTGAATTAATGGTGAACGGTACTGCCGACCTTC 2280
DB 2221 GCATGGTCCAAATCGACACACGCAAAAGTGAATTAATGGTGAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGTCACACCATCCAAACCGTTGTTGGCGTCTGAGCGAGGTATTAACGCCGCCAGT 2340
DB 2281 AAGGTCACACCATCCAAACCGTTGTTGGCGTCTGAGCGAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
DB 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCCGTCGGTCCGTCGCTGAACTCTTACGAGGAAGAG 2460
DB 2401 GAAGCGGTAAATAAGACAAACCCGTCGGTCCGTCGCTGAACTCTTACGAGGAAGAG 2460
QY 2461 TTGGCAAAAGATCCACGTTATTCGCCCAACCATGGAACACGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGGCAAAAGATCCACGTTATTCGCCCAACCATGGAACACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGAGATGTCGCTTTCTGCTGATGCGCTGCTGCTGCTGATCAACGCC 2580
DB 2521 CCGAACATCCCGAGATGTCGCTTTCTGCTGATGCGCTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
DB 2581 GCCAGCGTCTGACGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
QY 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
DB 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
QY 2701 GGATCCTCTCTCTGTCGGCCGAGCGCGCTTCGAGCGCGG 2741
DB 2701 GGATCCTCTCTCTGTCGGCCGAGCGCGCTTCGTCGGGAGG 2741

RESULT 7
AX172306
LOCUS AX172306 8101 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 5 from Patent WO0144467.
ACCESSION AX172306
VERSION AX172306.1 GI:14597487
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Goubin-Gramatica,F., Ducommun,B. and Prevost,G.
TITLE Method for obtaining human cdc25 phosphatases and method for
identifying human cdc25 phosphatase modulators
JOURNAL Patent: WO 0144467-A 5 21-JUN-2001;
SOCIETE DE CONSEILS DE RECHERCHES ET D'APPLICATIONS SCIENTIFIQUES

FEATURES		(FR)		Location/Qualifiers	
source		1..8101		/organism="synthetic construct"	
				/mol_type="unassigned DNA"	
				/db_xref="taxon:32630"	
				/note="souche JM109 / pMAL - Hs Cdc25C"	
ORIGIN					
Query Match		82.0%; Score 2705.8; DB 6; Length 8101;			
Best Local Similarity		99.9%; Pred. No. 0;			
Matches 2707;		Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATATAGAGCCCGGAAGAGA	60		
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATATAGAGCCCGGAAGAGA	60		
Qy	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGATATGCCG	120		
Db	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGATATGCCG	120		
Qy	121	GTGTCCTTATACAGACCGTTTCCCGCGTGTGAACCGAGCCAGCACGTTTCTCGGAAA	180		
Db	121	GTGTCCTTATACAGACCGTTTCCCGCGTGTGAACCGAGCCAGCACGTTTCTCGGAAA	180		
Qy	181	CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240		
Db	181	CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240		
Qy	241	AACAACTGGCGGCAACAGTCTGTGATGTGGGTGGCCACCTCGATGTGGCCCTGC	300		
Db	241	AACAACTGGCGGCAACAGTCTGTGATGTGGGTGGCCACCTCGATGTGGCCCTGC	300		
Qy	301	ACGCGCGTCCCAATTTGTCCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360		
Db	301	ACGCGCGTCCCAATTTGTCCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360		
Qy	361	TGGTGGTTCGATGGTGAAGCGCGCTGGAAGCGTGTGAAAGCGCGTGTGACCAATC	420		
Db	361	TGGTGGTTCGATGGTGAAGCGCGCTGGAAGCGTGTGAAAGCGCGTGTGACCAATC	420		
Qy	421	TTCTCGGCAACCGGTACAGTGGGTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAGT	480		
Db	421	TTCTCGGCAACCGGTACAGTGGGTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAGT	480		
Qy	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCGGCGTTATTTCTTGATGTCTCTGACCA	540		
Db	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCGGCGTTATTTCTTGATGTCTCTGACCA	540		
Qy	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGGTGGAGCATC	600		
Db	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGGTGGAGCATC	600		
Qy	601	TGGTGCATTTGGTTCACAGCAATCGCGTGTAGCGGCGCCATTAAAGTTCTGCTCGG	660		
Db	601	TGGTGCATTTGGTTCACAGCAATCGCGTGTAGCGGCGCCATTAAAGTTCTGCTCGG	660		
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
Qy	721	CGGAACGGGAAGGCACTGGAGTGCATATCTCGCGTTTCAACAAACCAATGCAATGTGA	780		
Db	721	CGGAACGGGAAGGCACTGGAGTGCATATCTCGCGTTTCAACAAACCAATGCAATGTGA	780		
Qy	781	ATGAGGCGATCGTTCCCACTTCGATGTGGTTCGAACGATCAGATGCGCTGGCGCAA	840		
Db	781	ATGAGGCGATCGTTCCCACTTCGATGTGGTTCGAACGATCAGATGCGCTGGCGCAA	840		
Qy	841	TGCGCGCATTCAGATTCGGGCTGCGGTGTGGATATCTCGGTAGTGGGATACG	900		
Db	841	TGCGCGCATTCAGCGAGTTCGGGCTGCGGTGTGGGATATCTCGGTAGTGGGATACG	900		

Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGGATTTTC	960		
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGGATTTTC	960		
Qy	961	GCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020		
Db	961	GCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020		
Qy	1021	AGGCAATCAGCTGTGCGCGCTCTCACTGCTGAAGAAAAACACCTGGCGCCCAATA	1080		
Db	1021	AGGCAATCAGCTGTGCGCGCTCTCACTGCTGAAGAAAAACACCTGGCGCCCAATA	1080		
Qy	1081	CGCAAAACCGGCTCTCCCGCGCGTTGGCCGATTCATTAAATGCAGCTGCACGACGTTT	1140		
Db	1081	CGCAAAACCGGCTCTCCCGCGCGTTGGCCGATTCATTAAATGCAGCTGCACGACGTTT	1140		
Qy	1141	CCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGCTCACTCAATTAG	1200		
Db	1141	CCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGCTCACTCAATTAG	1200		
Qy	1201	GCACAAATTCATGTTTGCAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260		
Db	1201	GCACAAATTCATGTTTGCAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260		
Qy	1261	TCAGGACAGCCATCGGAAGCTGTGTATGGCTGTGCGAGTGTAAATCACTGCATAATTTCG	1320		
Db	1261	TCAGGACAGCCATCGGAAGCTGTGTATGGCTGTGCGAGTGTAAATCACTGCATAATTTCG	1320		
Qy	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT	1380		
Db	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT	1380		
Qy	1381	CTGCAAAATTTCTGAAATCAGCTGTTTGACAAATTAATCATCGGCTCGTAAATGTGGGA	1440		
Db	1381	CTGCAAAATTTCTGAAATCAGCTGTTTGACAAATTAATCATCGGCTCGTAAATGTGGGA	1440		
Qy	1441	ATTGTGAGCGGATAACAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500		
Db	1441	ATTGTGAGCGGATAACAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500		
Qy	1501	GCACTTCCACCAACAGGACCATAGATTATGAAAATCTGAAGAAGTAAATCTGGTAATCTCG	1560		
Db	1501	GCACTTCCACCAACAGGACCATAGATTATGAAAATCTGAAGAAGTAAATCTGGTAATCTCG	1560		
Qy	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTAAAGAAATTCGAGAAGAT	1620		
Db	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTAAAGAAATTCGAGAAGAT	1620		
Qy	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT	1680		
Db	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT	1680		
Qy	1681	GGGCAACTGGCGATGGCGCTGACATATCTTCTGGGCACACCGCTTTGGTGGCTAC	1740		
Db	1681	GGGCAACTGGCGATGGCGCTGACATATCTTCTGGGCACACCGCTTTGGTGGCTAC	1740		
Qy	1741	GCTCAATCTGGCTGTGTTGGCTGAAATCACCCCGCAAAAGCGTTCAGGACAAAGCTGTAT	1800		
Db	1741	GCTCAATCTGGCTGTGTTGGCTGAAATCACCCCGCAAAAGCGTTCAGGACAAAGCTGTAT	1800		
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTAACAACGCAAGCTGATTGCTTACCCGATCGCTGTT	1860		
Db	1801	CCGTTTACCTGGGATGCGGTACGTTAACAACGCAAGCTGATTGCTTACCCGATCGCTGTT	1860		
Qy	1861	GAAGCGTTATCGCTGATTTTATAAAGATCTGCTGCCGAACCGCCCAAAAACCTGGAA	1920		
Db	1861	GAAGCGTTATCGCTGATTTTATAAAGATCTGCTGCCGAACCGCCCAAAAACCTGGAA	1920		
Qy	1921	GAGATCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGTGTTTCAAC	1980		
Db	1921	GAGATCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGTGTTTCAAC	1980		
Qy	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGCTGACGGGGTTATGCTTCAAG	2040		

|||||
Db 1 CCGACCAATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAACGTATTACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAACGTATTACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCCAGCCAGCGTTTCTCGGAAAA 180
Db 121 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCCAGCCAGCGTTTCTCGGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGGATGTCGGAGCTGAATTCACATCCCAACCGGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGATGTCGGAGCTGAATTCACATCCCAACCGGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTGGTTCGATTTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTGGTTCGATTTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 AGCGCGGTGCGCAAAATTTGTCGGCGGATTAATTCGCGCGGATCAACTGGGTGCCAGG 360
Db 301 AGCGCGGTGCGCAAAATTTGTCGGCGGATTAATTCGCGCGGATCAACTGGGTGCCAGG 360
QY 361 TGGTGGTCTCGATGCTAGAACGAGCGGCTGGAAGCGGCTGTAAGCGGCGGTGCACAATC 420
Db 361 TGGTGGTCTCGATGCTAGAACGAGCGGCTGGAAGCGGCTGTAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGGATGCCA 480
QY 481 TTGCTGTGGAAGTCCCTGCGACATAATGTTCCGGCGTATTTCTGATGCTCTGACACAGA 540
Db 481 TTGCTGTGGAAGTCCCTGCGACATAATGTTCCGGCGTATTTCTGATGCTCTGACACAGA 540
QY 541 CACCCATCAACAGTATATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTATAGCGGGCCCAATTAAGTTCGTCTCGG 660
Db 601 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTATAGCGGGCCCAATTAAGTTCGTCTCGG 660
QY 661 CGCGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTCCCACTGGCATGTGGTTGCCAAGCATGCGCTGGCGCGCA 840
Db 781 ATGAGGGCATCGTCCCACTGGCATGTGGTTGCCAAGCATGCGCTGGCGCGCA 840
QY 841 TGCGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGCGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
QY 961 GCCTGTGGGCAAAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGTGGGCAAAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGTGTGGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGTGTGGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGCACTGGCAACAGAGTTT 1140
|||||

Db 1081 CGCAAAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGCACTGGCAACAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGACGGTGCACCAATGCTTCTGGGG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGACCGTGCACCAATGCTTCTGGGG 1260
QY 1261 TCAGGACGCATTCGGAAGCTGTGGTGTGAGCTGTCAGGTGCTGAATCACTGTCATAAATTCG 1320
Db 1261 TCAGGACGCATTCGGAAGCTGTGGTGTGAGCTGTCAGGTGCTGAATCACTGTCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTCGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTCGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATATCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAATATCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
QY 1441 ATTTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCAACA 1500
Db 1441 ATTTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCAACA 1500
QY 1501 GCATTTCACCAACGAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTTGTAATCTGG 1560
Db 1501 GCATTTCACCAACGAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTTGTAATCTGG 1560
QY 1561 ATTTAAGCGGATTAAGGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGGATTAAGGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAATTCGCCACAGTT 1680
QY 1681 CGCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGTGGCTGAAATCAACCCGGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAAATCAACCCGGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTTGCTGCTGACGGGGTTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTTGCTGCTGACGGGGTTTATCGGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGTGACTGATTAATAACAAACATGAAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACTGATTAATAACAAACATGAAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTCTTAAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTCTTAAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
|||||


```
QY 2221 GCATGTCACATCGACACCGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGTCACATCGACACCGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGATTCTCGAAAACTATCTGCTGACTGATGAAGGTCGT 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGATTCTCGAAAACTATCTGCTGACTGATGAAGGTCGT 2400
QY 2401 GAAGCGTTATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTAGCGAAGAG 2460
Db 2401 GAAGCGTTATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTAGCGAAGAG 2460
QY 2461 TTGGCGAAGATCCACGATTGTCGCCCAACCATGGAAGAGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGATTGTCGCCCAACCATGGAAGAGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGGTGCGTACTGCGGTGATCAAGGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGGTGCGTACTGCGGTGATCAAGGCC 2580
QY 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAGAGCGGCGAGACTAAATTCAGGCTCG 2640
Db 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAGAGCGGCGAGACTAAATTCAGGCTCG 2640
QY 2641 AACAAACAACAATAAACAATAACAACCACTCGGATCGAGGGAAGGATTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAACAATAACAACCACTCGGATCGAGGGAAGGATTTCAGAAATTC 2700

RESULT 10
AR031992
LOCUS AR031992 7475 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5866684.
ACCESSION AR031992
VERSION AR031992.1 GI:5946281
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 7475)
AUTHORS Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Willson,F.Xavier.
TITLE Peptidyl inhibitors of viral proteases
JOURNAL Patent: US 5866684-A 1 02-FEB-1999;
FEATURES
Source Location/Qualifiers
1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCGAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTCGAGATATGCCG 120
Db 61 GTCGAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTCGAGATATGCCG 120
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGTGAACCAAGGCGACCGCTTTCTCGGAAA 180
Db 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGTGAACCAAGGCGACCGCTTTCTCGGAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC
```

```
Db 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACTCTCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACTCTCAGTCTGGCCCTGC 300
QY 301 ACGCCCGTTCGCAAAATTTGTCGGCGGANTAAATCTCCGGCGGATCAACTGGGTGCGCAGC 360
Db 301 ACGCCCGTTCGCAAAATTTGTCGGCGGANTAAATCTCCGGCGGATCAACTGGGTGCGCAGC 360
QY 361 TGTGTTGTTTCSATGGTGAAGACGAAGCGCGCTCGAAGCTCTGTAAGCGCGCTGCACAATC 420
Db 361 TGTGTTGTTTCSATGGTGAAGACGAAGCGCGCTCGAAGCTCTGTAAGCGCGCTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTAATTAATTAATTAATTAATTAATTA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTAATTAATTAATTAATTAATTAATTA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAGA 540
QY 541 CACCAATCAACAGTATTAATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCAATCAACAGTATTAATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGTGCGCATTTGGGTCACCGCAATCGCGCTGTTAGCGGCGCCATTAAGTTCTGTCTCGG 660
Db 601 TGTGCGCATTTGGGTCACCGCAATCGCGCTGTTAGCGGCGCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCGTAAATATCTCACTCGCAATCAAAATTCAGCCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCGTAAATATCTCACTCGCAATCAAAATTCAGCCCGATAG 720
QY 721 CGGAAACGGGAAGCGGCTGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGCGGCTGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCGATCGTTCCCGCTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGCGATCGTTCCCGCTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGC CGCGCCATTACCGAGTCCGGCTCGGGCTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGC CGCGCCATTACCGAGTCCGGCTCGGGCTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGTTAAACCCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGTTAAACCCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCGAAACCGAGCTGGAGCGCTTGTGCAACTCTCTCAGGGCGGCGGTGA 1020
Db 961 GCCTGCTGGGCGAAACCGAGCTGGAGCGCTTGTGCAACTCTCTCAGGGCGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCGCGTCTCACTGTTGAAAGAAAAACCCACCTGGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCGCGTCTCACTGTTGAAAGAAAAACCCACCTGGGCGCCCAATA 1080
QY 1081 CGCAACCCGCTCTCCCGCGGTTGGCGGATTCATTATGAGCTGGGCGACGACGAGTTT 1140
Db 1081 CGCAACCCGCTCTCCCGCGGTTGGCGGATTCATTATGAGCTGGGCGACGACGAGTTT 1140
QY 1141 CCCGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTCAGTTAGTCTACTCATTTAG 1200
Db 1141 CCCGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTCAGTTAGTCTACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTGACAGCTTATCATCGACTGCGACCGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTCATGTTGACAGCTTATCATCGACTGCGACCGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGAGCATCGGAAGCTGTTGGTATGCTGTCAGGTCGTAAATCACTGCAATATTCG 1320
Db 1261 TCAGGCGAGCATCGGAAGCTGTTGGTATGCTGTCAGGTCGTAAATCACTGCAATATTCG 1320
```

1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
1381 CTGGCAAAATATTCTGAAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTTGGA 1440
1381 CTGGCAAAATATTCTGAAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTTGGA 1440
1441 ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTTAGGTGTTTTCAGGA 1500
1441 ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTTAGGTGTTTTCAGGA 1500
1501 GCATCTCAACACAGGACCATAGATTATGAACCTGAAGAGTAAACTGGTAACTCTGG 1560
1501 GCATCTCAACACAGGACCATAGATTATGAACCTGAAGAGTAAACTGGTAACTCTGG 1560
1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
1741 GCTCAATCTGGCCTGTGCTGAAATCACCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
1741 GCTCAATCTGGCCTGTGCTGAAATCACCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
1801 CCGTTTACCTGGGATGCGGTACGTTTACAAAGCAAGCTGATGCTTACCCGATCGCTGT 1860
1801 CCGTTTACCTGGGATGCGGTACGTTTACAAAGCAAGCTGATGCTTACCCGATCGCTGT 1860
1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTGGCGAACCCGCCCAAAAACCTTGGGAA 1920
1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTGGCGAACCCGCCCAAAAACCTTGGGAA 1920
1921 GAGATCCCGGCGCTCGATTAAGAACTGAAGGAAAGGTAAGAGCGCGTGAATTTCAAC 1980
1921 GAGATCCCGGCGCTCGATTAAGAACTGAAGGAAAGGTAAGAGCGCGTGAATTTCAAC 1980
1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGGGTGGATAACGCTGGCGCGAAAGCG 2100
2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGGGTGGATAACGCTGGCGCGAAAGCG 2100
2101 GGTCTGACCTTCTGTTGACCTGATTAAGAAACACATCAATGAATGCGACACCGATTAAC 2160
2101 GGTCTGACCTTCTGTTGACCTGATTAAGAAACACATCAATGAATGCGACACCGATTAAC 2160
2161 TCCATCGCAAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGGCCGTGG 2220
2161 TCCATCGCAAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGGCCGTGG 2220
2221 GCATGGTCCAAATCGACACAGCAAGAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
2221 GCATGGTCCAAATCGACACAGCAAGAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
2281 AAGGTCACACCATCCAAACCGTTTCTGCGGCTGCTGAGCGAGGATTAACCGCCCGCAGT 2340
2281 AAGGTCACACCATCCAAACCGTTTCTGCGGCTGCTGAGCGAGGATTAACCGCCCGCAGT 2340
2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400

2401 GAAGCGTTAATAAGACAAACCCCTGGTGGCTAGCGCTGAAGTCTTTACGAGGAAG 2460
2401 GAAGCGTTAATAAGACAAACCCCTGGTGGCTAGCGCTGAAGTCTTTACGAGGAAG 2460
2461 TTGCGGAAAGATCCACGCTATTGCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
2461 TTGCGGAAAGATCCACGCTATTGCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
2521 CCGAACATCCGCGAGATGTCGGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
2521 CCGAACATCCGCGAGATGTCGGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGGAAGACGCGAGACTAAATTCGAGCTCG 2640
2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGGAAGACGCGAGACTAAATTCGAGCTCG 2640
2641 AACAAACAAACAATAACAATAACAACACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
2641 AACAAACAAACAATAACAATAACAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700

RESULT 11
AR207294
LOCUS AR207294 7475 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6372883.
ACCESSION AR207294
VERSION AR207294.1 GI:21506162
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,F.Xavier.
TITLE Antiviral medicaments
JOURNAL Patent: US 6372883-A 1 16-APR-2002;
FEATURES Location/Qualifiers
source
1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAAAGTATATGATGTCGAGAGTATCGG 120
DB 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAAAGTATATGATGTCGAGAGTATCGG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGTAAAGTATGATGTCGAGAGTATCGG 180
DB 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGTAAAGTATGATGTCGAGAGTATCGG 180
QY 181 CGCGGAAAAGTGGAGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
DB 181 CGCGGAAAAGTGGAGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
QY 241 AACAACTGGCGGCAACAGCTGTTGCTGATTTGGCGTTGCCACCTTCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAACAGCTGTTGCTGATTTGGCGTTGCCACCTTCAGTCTGGCCCTGC 300
QY 301 AGCGCCCTCGCAAAATTTGTCGGGGGATTAATCTCGCGCGCATCAACTGGGTGCGAGG 360
DB 301 AGCGCCCTCGCAAAATTTGTCGGGGGATTAATCTCGCGCGCATCAACTGGGTGCGAGG 360
QY 361 TGGTGGTCTCGATGTAAGCAAGCGCGCTGAGAGCTGTAAAGCGCGGTGCAACAATC 420
DB 361 TGGTGGTCTCGATGTAAGCAAGCGCGCTGAGAGCTGTAAAGCGCGGTGCAACAATC 420

QY 421 TTCTCGCGCAACGCGTCAGTGGCTGATCATTAATACTATCCGCTGGATGACAGGATGCCA 480
DB |||||
QY 421 TTCTCGCGCAACGCGTCAGTGGCTGATCATTAATACTATCCGCTGGATGACAGGATGCCA 480
DB |||||
QY 481 TTGCTGTGGAGCTGCCCTGCCTAATAGTTCGCGCTTAATTTCTTGATGCTCTGACACAGA 540
DB |||||
QY 481 TTGCTGTGGAGCTGCCCTGCCTAATAGTTCGCGCTTAATTTCTTGATGCTCTGACACAGA 540
DB |||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCACTGGGCGGTGGAGCATC 600
DB |||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCACTGGGCGGTGGAGCATC 600
DB |||||
QY 601 TGGTCGCATTTGGTCCACAGCAAAATCGGCTGTTAGCGGGGCCCAATTAAGTTCTGTCTCGG 660
DB |||||
QY 601 TGGTCGCATTTGGTCCACAGCAAAATCGGCTGTTAGCGGGGCCCAATTAAGTTCTGTCTCGG 660
DB |||||
QY 661 CGCGTCTCGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
DB |||||
QY 661 CGCGTCTCGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
DB |||||
QY 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB |||||
QY 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB |||||
QY 781 ATGAGGGCATCGTTCCTCACTCGATGCTGGTTGCCAACGATGCGCGCTGGCGCGCA 840
DB |||||
QY 781 ATGAGGGCATCGTTCCTCACTCGATGCTGGTTGCCAACGATGCGCGCTGGCGCGCA 840
DB |||||
QY 841 TGGCGCCATTAACGAGTCCGGCTGCGCTGGTGGTGGGATATCTCGTAGTGGGATACG 900
DB |||||
QY 841 TGGCGCCATTAACGAGTCCGGCTGCGCTGGTGGTGGGATATCTCGTAGTGGGATACG 900
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
DB |||||
QY 961 GCCTGTGGGGCAACCAAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB |||||
QY 961 GCCTGTGGGGCAACCAAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB |||||
QY 1021 AGGCAATACAGCTGTGCGCGCTCACTGTGTAAGAAAGAAACCAACCTGGCGCGCAATA 1080
DB |||||
QY 1021 AGGCAATACAGCTGTGCGCGCTCACTGTGTAAGAAAGAAACCAACCTGGCGCGCAATA 1080
DB |||||
QY 1081 GCGAAAACGCGCTCTCCGCGCGCTTGGCGGATTCATTAATGACGCTGGCAGCAGGTTT 1140
DB |||||
QY 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG 1200
DB |||||
QY 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG 1200
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACGTGCACGCTGCACCAATGCTTCTGGCG 1260
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACGTGCACGCTGCACCAATGCTTCTGGCG 1260
DB |||||
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCAAGTCACTGCAATAATTCG 1320
DB |||||
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCAAGTCACTGCAATAATTCG 1320
DB |||||
QY 1321 TGTGCTCAAGCGGCACCTCCGTTCTCGGTAATGTTTTTTCGCGCGGACATCAATAACGGTT 1380
DB |||||
QY 1321 TGTGCTCAAGCGGCACCTCCGTTCTCGGTAATGTTTTTTCGCGCGGACATCAATAACGGTT 1380
DB |||||
QY 1381 CTGGCAAAATTTCTGAATAGCTGTTGACAAATTAATCACTCGGCTCGTATAATGTTGGA 1440
DB |||||
QY 1381 CTGGCAAAATTTCTGAATAGCTGTTGACAAATTAATCACTCGGCTCGTATAATGTTGGA 1440
DB |||||
QY 1441 ATTGTGAGCGGATCAAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
DB |||||
QY 1441 ATTGTGAGCGGATCAAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
DB |||||

QY 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAACTCGG 1560
DB |||||
QY 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAACTCGG 1560
DB |||||
QY 1561 ATTAACCGCGGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1561 ATTAACCGCGGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCACAGGTT 1680
DB |||||
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCACAGGTT 1680
DB |||||
QY 1681 GCGCAACTCGCGATGCGCCCTGCATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB |||||
QY 1681 GCGCAACTCGCGATGCGCCCTGCATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB |||||
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
DB |||||
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
DB |||||
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTCCTACCGGATCGCTGTT 1860
DB |||||
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTCCTACCGGATCGCTGTT 1860
DB |||||
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTCTGCCGAACCGCCCAAAACCTCGGAA 1920
DB |||||
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTCTGCCGAACCGCCCAAAACCTCGGAA 1920
DB |||||
QY 1921 GAGATCCGCGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGTATGTTCAAC 1980
DB |||||
QY 1921 GAGATCCGCGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGTATGTTCAAC 1980
DB |||||
QY 1981 CTCGAAGAACCGTACCTTACCTGCGCGCTGATTGCTGACCGGGGTTATGGTTCAAG 2040
DB |||||
QY 1981 CTCGAAGAACCGTACCTTACCTGCGCGCTGATTGCTGACCGGGGTTATGGTTCAAG 2040
DB |||||
QY 2041 TATGAAAAACGCGCAAGTACGACATTAAGAGCGTGGGCTGGATAACGCTGGCGCGAAAGCG 2100
DB |||||
QY 2041 TATGAAAAACGCGCAAGTACGACATTAAGAGCGTGGGCTGGATAACGCTGGCGCGAAAGCG 2100
DB |||||
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAAACATGAAATGCGAGACACCGATTAC 2160
DB |||||
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAAACATGAAATGCGAGACACCGATTAC 2160
DB |||||
QY 2161 TCCATCGCAAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCGGTGG 2220
DB |||||
QY 2161 TCCATCGCAAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCGGTGG 2220
DB |||||
QY 2221 GCATGTCCAAACATCGACACACCGCAAGTAAATTTATGGTGTAAACGCTACTGCCGACCTTC 2280
DB |||||
QY 2221 GCATGTCCAAACATCGACACACCGCAAGTAAATTTATGGTGTAAACGCTACTGCCGACCTTC 2280
DB |||||
QY 2281 AAGGTCACACCTCCAAACCGTTTGGGCTGCTGAGCGCAGGTATTAAACGCGCGCAGT 2340
DB |||||
QY 2281 AAGGTCACACCTCCAAACCGTTTGGGCTGCTGAGCGCAGGTATTAAACGCGCGCAGT 2340
DB |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
DB |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
DB |||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAGAG 2460
DB |||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAGAG 2460
DB |||||
QY 2461 TTGGCCGAAAGATCCAGTATTCGCCACCATGGAACAAACCGCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2461 TTGGCCGAAAGATCCAGTATTCGCCACCATGGAACAAACCGCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2521 CCGAAACATCCCGCAGATGTCGCGTTTCTGGTATGCGGTGCGTACTCGCGTGTATCAACGCC 2580
DB |||||
QY 2521 CCGAAACATCCCGCAGATGTCGCGTTTCTGGTATGCGGTGCGTACTCGCGTGTATCAACGCC 2580
DB |||||
QY 2581 GCCAGCGGTGCTCAGACTGTCGATGAGGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG 2640
DB |||||

Db 2581 GCCAGGGTGTGACACTGTGATGAGCCCTGAAGAAGCGGCAGACTTAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
RESULT 12
AX377532
LOCUS AX377532 9191 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 9 from Patent WO0212553.
ACCESSION AX377532
VERSION AX377532.1 GI:19573718
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdoerf,H. and Muth,J.
TITLE Method for detecting mutations in nucleotide sequences
JOURNAL Patent: WO 0212553-A 9 14-FEB-2002;
Nanogen Recognomics GmbH (DE)
FEATURES
source Location/Qualifiers
1..9191
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pMALc2x-mutS Plasmid"
ORIGIN
Query Match 81.8%; Score 2698; DB 6; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCACAAACCTTTCCGGGTATGGCATGATAGGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCACAAACCTTTCCGGGTATGGCATGATAGGCCCGGAAGAGA 60
Qy 61 GTCATTTGAGGTGGTGAATGTGAACACAGTAACGTTATACCATGTCGCGAGATATGCG 120
Db 61 GTCATTTGAGGTGGTGAATGTGAACACAGTAACGTTATACCATGTCGCGAGATATGCG 120
Qy 121 GTGCTCTTATCAGACCGGTTTCCCGGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAA 180
Db 121 GTGCTCTTATCAGACCGGTTTCCCGGTGGTGAACACAGCCAGCCAGTTTCTGCGAAAA 180
Qy 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTCCTCAACCCGCTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTCCTCAACCCGCTGGCAC 240
Qy 241 AACAACTGCGGGCAACAGCTGTTGCTGATGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGCAACAGCTGTTGCTGATGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTTCGCAAAATTTGTCGGGGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTTGTCGGGGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGGTTCGATGTTAGAACGAAGCGGCTGAAAGCCCTGTAAGCGGCGGTGCAATTC 420
Db 361 TGGTGGTTCGATGTTAGAACGAAGCGGCTGAAAGCCCTGTAAGCGGCGGTGCAATTC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCTCCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCTCCGCTGGATGACAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Qy 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGAAGCGGTACCGCATCTGGCGGTGGAGCATC 600

Db 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGAAGCGGTACCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTTCGCAATTCGGGTACACAGCAAAATCGCGCTGTTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTTCGCAATTCGGGTACACAGCAAAATCGCGCTGTTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTCGGCTCGGCTGGCGATTAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGGCTCGGCTGGCGATTAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCGACTGGAAGTGCATGTCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAAGTGCATGTCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCAACGATCAGATGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCAACGATCAGATGCGCTGGGCGCAA 840
Qy 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACCATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAAAACACAGCGTGAACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACACAGCGTGAACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGTTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGTTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCACGACAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCACGACAGGTTT 1140
Qy 1141 CCCGACTCGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTTAGTCTCACTCATTTAG 1200
Db 1141 CCCGACTCGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTTAGTCTCACTCATTTAG 1200
Qy 1201 GCACAATTTCTATGTTGACAGCTTATCATCGATGCAACGGTGCAACCAATGTTCTTGGCG 1260
Db 1201 GCACAATTTCTATGTTGACAGCTTATCATCGATGCAACGGTGCAACCAATGTTCTTGGCG 1260
Qy 1261 TCAGGCGCCATCGGAAGCTGTGGTATGCTGTCAGGTCGTAATCACTGCAATATTCG 1320
Db 1261 TCAGGCGCCATCGGAAGCTGTGGTATGCTGTCAGGTCGTAATCACTGCAATATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACCTCCCGTTCGATTAATGTTTTTGGCGCGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCCGTTCGATTAATGTTTTTGGCGCGACATCATACCGGTT 1380
Qy 1381 CTGGCAAAATATTTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATATTTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGGTTGTTTTCAGA 1500
Db 1441 ATTGTGAGCGGATACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGGTTGTTTTCAGA 1500
Qy 1501 GCACCTTCCAAACAGGACCATAGCATATGAAAATCGAAGAAGGTTAACTGGTATCTCG 1560
Db 1501 GCACCTTCCAAACAGGACCATAGCATATGAAAATCGAAGAAGGTTAACTGGTATCTCG 1560
Qy 1561 ATTTAACCGCGATTAAGGCTATAA CCGTCTCGCTGAAGTCCGGTGAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAACCGCGATTAAGGCTATAA CCGTCTCGCTGAAGTCCGGTGAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTTCCACAGGTT 1680

Db 1621 ACGGAATTAAAGTCAACCGTTGAGCATCCGGATAAACTCGAAGAGAAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGCGCCTGACATTATCTCTGGGCACACAGCCGCTTTCGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGCGCCTGACATTATCTCTGGGCACACAGCCGCTTTCGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGACAAAGCGTTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGACAAAGCGTTTCAGGACAAGCTGTAT 1800
Qy 1801 CGTTTACCTGGATGCGCTAGCTTACACGCAAGCTGATGCTTACCCGATCCCTGTT 1860
Db 1801 CGTTTACCTGGATGCGCTAGCTTACACGCAAGCTGATGCTTACCCGATCCCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTCGCGCTGATTGCTGCTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTCGCGCTGATTGCTGCTGACGGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGAACCTGGCGCTGGATACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGAACCTGGCGCTGGATACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTAAGAACCAACACATGATGACGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAAGAACCAACACATGATGACGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCCCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATATGTTGTAAACGCTGACGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATATGTTGTAAACGCTGACGACCTTC 2280
Qy 2281 AAGGTCACCATCAACACCGTTCTGTTGGCTGCTGACGCGAGGTATTAACGGCCGCTCAGT 2340
Db 2281 AAGGTCACCATCAACACCGTTCTGTTGGCTGCTGACGCGAGGTATTAACGGCCGCTCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTCGT 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTCGT 2400
Qy 2401 GAAGCGGTTAATAAGCAAAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGCAAAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGCGGAAGATCCACGCTATTGCGCGCACCATGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAGATCCACGCTATTGCGCGCACCATGAAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGCTTTCTGTTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTTCTGTTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGCCGCGACACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGCCGCGACACTAATTCGAGCTCG 2640
Qy 2641 AACCAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCC 2706
Db 2701 GGATCC 2706

RESULT 13
AF097412
LOCUS
DEFINITION
AF097412
ACCESSION
AF097412.1 GI:3983120
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS
promoter
gene
CDS
misc_feature
misc_feature
misc_feature

AF097412 6724 bp DNA circular SYN 28-JAN-2002
Expression vector pMBP-parallel1, complete sequence.
AF097412
AF097412.1 GI:3983120
Expression vector pMBP-parallel1
Expression vector pMBP-parallel1
other sequences; attficial sequences; vectors.
1 (bases 1 to 6724)
Sheffield, P., Garrard, S. and Derewenda, Z.
Overcoming expression and purification problems of RhoGDI using a
family of 'parallel' expression vectors
Protein Expr. Purif. 15 (1), 34-39 (1999)
99150479
10024467
2 (bases 1 to 6724)
Sheffield, P.J., Garrard, S.M. and Derewenda, Z.S.
Direct Submission
Submitted (05-OCT-1998) Molecular Physiology and Biological
Physics, University of Virginia, 4215 Jordan Hall, 1300 Jefferson
Park Avenue, Charlottesville, Virginia 22908, USA
Location/Qualifiers
1..6724
/organism="Expression vector pMBP-parallel1"
/mol_type="other DNA"
/db_xref="taxon:84479"
204..1163
/gene="lacIq"
204..1163
/gene="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAC83812.1"
/db_xref="GI:3983121"
/translation="MAELNYIPNRVAQLAGKQSLIGVATSLALHAPSOIVAAIKS
RADQLGASVVSVSVERSGVACKAAVHNLQAQVSGLIINYPDDODDAIVAAACATNV
PAFLDVSQTPINSIIFSHEDTGRVLAHQVGLIINYPDDODDAIVAAACATNV
SKYLTRNQIQAEREGDWSAMSGFOQTQMLNEGIVPTAMLVANDQMALGAMRAITE
HGLRVGADISVYGGDDTSSCVIPPLTTIKODFRLLGQTSVDRLLQLSQGQAVKGNQ
LLPVSLVRKKTLLAPNTQTASPRALADSLMQLARQVRSLESG"
1453..1481
/note="tac; inducible promoter"
1528..2294
/gene="male"
1528..2294
/gene="male"
/codon_start=1
/transl_table=11
/product="maltose binding protein"
/protein_id="AAC83813.1"
/db_xref="GI:3983122"
/translation="MKIEGKLVIVINGDKYNGLAEVGKPEKDTGKIVTVEHPDKL
EKKFQVAATGDPDIIIFWAHDFGQYAGSLAEITPDKAQDKLIPFTWDAVRVNG
KLIAVPIAVEALSLIVNKDLLPNPKTWEEIIPALDKELKAGKSLMNLQEPYFWP
LTAADGGVAFYKNGKYDIKQGVNDAGAKGLTFLVDLKNKHNADTDYIAEAA
NKGETAMTINGPWAWSNIDTSKVNTGYTVLPTFGQGPSPKFPVGLSAGINAAAPKEL
AKEFLYENLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIATAATMENAQKEIMPNI
PQMSAFVAARVAVINAASGQTVDEALKAQNTSSNNNNNNNNNNLIGIDTNTLYF
QGAMPPEFKGLRRRAQLVRPLSNLEPAVSRHAPVSLALAVLQRRDWNENPGVTQLNRL
AAHPFPASWRNSEARTRDPSQQLRSLNGEMQLGCGFG"
2633..2688
/gene="male"
/note="protein spacer region"
2689..2709
/gene="male"
/note="Tobacco etch virus protease cleavage site"
2711..2810
/gene="male"

gene	/note="multiple cloning site"				
promoter	3426..4356	/gene="bla"			
	3426..3454	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
CDS	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
	3496..4356	/gene="bla"			
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				
3496..4356	/gene="bla"				

Db	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAAAATGCTGA	780	
Qy	781	ATGAGGGCATCGTTCCCACTCGCATGTGTTGCCAACGATCAGATGGCGTGGGCGCAA	840	
Db	781	ATGAGGGCATCGTTCCCACTCGCATGTGTTGCCAACGATCAGATGGCGTGGGCGCAA	840	
Qy	841	TGCGGCCATTACCGAGTCCGGCTCGCGTGGTGGGATATCTCGTAGTGGGATAG	900	
Db	841	TGCGGCCATTACCGAGTCCGGCTCGCGTGGTGGGATATCTCGTAGTGGGATAG	900	
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACACCATCAAAACAGGATTTTC	960	
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACACCATCAAAACAGGATTTTC	960	
Qy	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020	
Db	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020	
Qy	1021	AGGGCAATCAGCTGTTGCCCGTCTCAGCTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080	
Db	1021	AGGGCAATCAGCTGTTGCCCGTCTCAGCTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080	
Qy	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCACACAGGTTT	1140	
Db	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCACACAGGTTT	1140	
Qy	1141	CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200	
Db	1141	CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200	
Qy	1201	GCACAAATTCATGTTTGACAGCTTATCATCGATGCGGTGCACCAATGCTTCTGGCG	1260	
Db	1201	GCACAAATTCATGTTTGACAGCTTATCATCGATGCGGTGCACCAATGCTTCTGGCG	1260	
Qy	1261	TCAGCAGCCATCGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGCATTAATTCG	1320	
Db	1261	TCAGCAGCCATCGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGCATTAATTCG	1320	
Qy	1321	TGTGCTCAAGCGGCACCTCCCGTCTCGGATAATGTTTTTGGCGCGACATCAATACGGTT	1380	
Db	1321	TGTGCTCAAGCGGCACCTCCCGTCTCGGATAATGTTTTTGGCGCGACATCAATACGGTT	1380	
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440	
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440	
Qy	1441	ATTGTAGCGGATTAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500	
Db	1441	ATTGTAGCGGATTAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500	
Qy	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGTTAATCTGG	1560	
Db	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGTTAATCTGG	1560	
Qy	1561	ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620	
Db	1561	ATTAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620	
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAAAATCCACAGGTT	1680	
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAAAATCCACAGGTT	1680	
Qy	1681	GGGGCACTGGCGATGGCCCTGACATTAATCTTGGGCAACAGCCGCTTTGGTGGCTAC	1740	
Db	1681	GGGGCACTGGCGATGGCCCTGACATTAATCTTGGGCAACAGCCGCTTTGGTGGCTAC	1740	
Qy	1741	GCTCAATCTGGCTGTGGCTGAAATCAATCCCGGCAAAAGCGTTCCAGACAAGCTGTAT	1800	
Db	1741	GCTCAATCTGGCTGTGGCTGAAATCAATCCCGGCAAAAGCGTTCCAGACAAGCTGTAT	1800	
Qy	1801	CGTTTACCTGGGATGGCTAGCTTACAAACGGCAAGCTGATGCTTACCCGATCCCTGTT	1860	

Db 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATATCCGCGGTTAAACCAACCATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCACGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCACGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCCCTGGCGGCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCCCTGGCGGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGGCAACAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGGCAACAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCAATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCAATTAG 1200
Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGTTCGTAATCACTGCAATATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGTTCGTAATCACTGCAATATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAATATTTCTCAATAGCTGTGCACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATTTCTCAATAGCTGTGCACATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTAGGTGTTTTCACGA 1500
Qy 1501 GCATTTCACCAAGGACCATAGATTATGAAAACTGAAGAAAGTAAACTGTTGTAATCTGG 1560
Db 1501 GCATTTCACCAAGGACCATAGATTATGAAAACTGAAGAAAGTAAACTGTTGTAATCTGG 1560
Qy 1561 ATTAACGGGATAAAGGCTATAACGCTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATAACGCTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATCCCAAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATCCCAAGGTT 1680
Qy 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCGCTGTGGCTGAATCAACCCGGCAAGCGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCGCTGTGGCTGAATCAACCCGGCAAGCGTTCCAGGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCCGTAGCTTACACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCCGTAGCTTACACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTCCGAAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980

Qy 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGANTGCTGTAACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGANTGCTGTAACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCCTGGTTGACCTGATTTAAAAAACAACACATGAATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCCTGGTTGACCTGATTTAAAAAACAACACATGAATGACAGACCGGATTAC 2160
Qy 2161 TCCATCGCAAGCTGCTTCTTAAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAAGCTGCTTCTTAAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAACTCGACACCAAGAGTGAATTAATGTTGTAACGGTAACTCCGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACCAAGAGTGAATTAATGTTGTAACGGTAACTCCGACCTTC 2280
Qy 2281 AAGGTTCAACCATCCAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGTGGTCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGTCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGCGGAAAGATCCACGCTATTGCGGCCACCACTGAAAAACGCCAGAAAGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGCTATTGCGGCCACCACTGAAAAACGCCAGAAAGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGCGTTTCTGTTGATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGCGTTTCTGTTGATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGCTGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640

RESULT 15
AF031088 6706 bp DNA circular SYN 07-MAY-2001
LOCUS Shuttle vector pMAL-pIII, complete sequence.
DEFINITION AF031088
ACCESSION AF031088
VERSION AF031088.1 GI:2623821
KEYWORDS Shuttle vector pMAL-pIII
SOURCE Shuttle vector pMAL-pIII
ORGANISM Shuttle vector pMAL-pIII
OTHER SEQUENCES; ARTIFICIAL SEQUENCES; VECTORS.
REFERENCE 1 (bases 1 to 6706)
AUTHORS Zwick,M.B., Bonnycastle,L.L., Noren,K.A., Venturini,S., Leong,E.,
Barbas,C.F. III, Noren,C.J. and Scott,J.K.
TITLE The maltose-binding protein as a scaffold for monovalent display of
peptides derived from phage libraries
JOURNAL Anal. Biochem. 264 (1), 87-97 (1998)
MEDLINE 99002881
PUBMED 9784192
REFERENCE 2 (bases 1 to 6706)
AUTHORS Noren,K.A. and Noren,C.J.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road,
Beverly, MA 01915, USA
FEATURES
source
1..6706 Location/Qualifiers
/organism="Shuttle vector pMAL-pIII"
/mol_type="genomic DNA"

```

/db xref="taxon:69327"
/notes="derived from commercially available vector pMal-p2
(New England Biolabs)"
15..1163
/genes="lacI"
/promoter
/notes="E. coli gene"
15..46
/genes="lacI"
81..1163
/genes="lacI"
/notes="transcriptional repressor"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAB86558.1"
/db_xref="GI:2623822"
/translation="MKPVTLDYAEYAGYSYQTSRVNVQASHVSAKTRKVEAMAE
LNYIPNVRAQLAGKQSLIGVATSSLAHAPSQIVAAIKRADQLGASVVSVMSERS
GVBECKAAVHNLAAQVSGLIINYPLDDQDAIAVEAACTNVPALFLDYSDQTPINSII
FSDGDTRLGVHVALGHQIALLAGPLSSVSARLRAGWHKYLTRNQIOPIAREG
DWSAMGFOQTQMLNEGIVPTAMLVANDQMALGAMRAITBSGLRVGADISVVGVDIT
EDSSCVIPPLTTIKODPRLLGQTSVDRLLQLSQGAVKGNQLLPVSLVVKRTTLAPNT
QTASPRALADSLMQLARQVSRLESQ"
1406..1433
/promoter
/notes="Pac promoter"
1528..2976
/genes="male-lacZa fusion"
/notes="E. coli gene fusion"
1528..2976
/genes="malB-lacZa fusion"
/notes="MBP-lacZa; MBP modified to include phage M13
secretion signal sequence"
/codon_start=1
/transl_table=11
/product="maltose binding protein-lacZ alpha peptide
fusion protein precursor"
/protein_id="AAB86559.1"
/db_xref="GI:2623823"
/translation="MKKLLFAIPLVPPYSHSADMKIEEGKLIWIINDKGYGLAEV
KKAFKDTGIVTVHEPDKLEKFPQVAATGDGPDIIIFWAHDFRGVAGSGLLAETP
DKAFQDKLPTTDAVRNKGKLIAYPIAVEALSIIYNKDLLPNPKTWBEI PALDKEL
KAKGSLAMFNLQPPYFWPLIAADGGVAPKYGVKVDIKDVGVDNAGAKAGLFLVD
LKNKHMADTDYSIAEAPFNKGETAMTINGPMAWSNIDTSKNYGVTVLPFGQPS
KPFQVLSAGNAGASPNKELAKELENYLLTDEGLEAVNKDKPLGVALKSYEEELAK
DPRLTAATHNAQKEINPNIQMSAFWAVTAVINAASGRQTVDEALDKDAQTNSSN
NNNNNNLIGIERISFEFGSSRVDLQASLALAVLQLRDWNENPGVTQLNRLAAHPFP
ASWRNSEARTDRPSQQLRSLSNGEWQLGCFG"
1528..1581
/genes="male-lacZa fusion"
/notes="M13 secretion signal sequence"
1582..2973
/genes="malB-lacZa fusion"
/product="maltose binding protein-lacZ alpha peptide
fusion protein"
3478..4338
/genes="bla"
/notes="R-plasmid ampicillin resistance gene"
3478..4338
/genes="bla"
/notes="penicillinase"
/codon_start=1
/transl_table=11
/product="RTEM-1 beta-lactamase"
/protein_id="AAB86560.1"
/db_xref="GI:2623824"
/translation="MSIQHPRVRLIIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IEDLSNGKILSEFRPEERPMSTFKVLLCGAVLSRVDAQOQLGRRIHISQNDLVE
YSPVTEKHTLDGTVMTRELCSAAITMSDNTAANLLTIIGPKELTAFLNMGDHTRL
DRWPELNEAIPINDERDTTNPVAMATTLRKLLTGELLTLASRQLLDWMDKVGRL
LRGALPAGWFIADKSGAGERSGRIIALGPDGPKSRIVVIYITGSQATMDERNROIA
EIGASLIKHW"
4381..4893
/standard_name="M13 origin (+ -)"

```

```

ORIGIN
Query Match      79.7%; Score 2631; DB 12; Length 6706;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 60; Gaps 1;

QY 1 CGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
DB 1 CGGACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGGTGTGAATGTGAACACAGTAAAGCTTATACGATGTGCGAGAGTATGCCG 120
DB 61 GTCAAATTCAGGGTGTGAATGTGAACACAGTAAAGCTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGGTTTCCCGCGTGTGAACACAGGCGAGCCAGCTTCTGCGGAAA 180
DB 121 GTGCTCTTATCAGACCGGTTTCCCGCGTGTGAACACAGGCGAGCCAGCTTCTGCGGAAA 180
QY 181 CGCGGGAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
DB 181 CGCGGGAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGCGCGTTGCCACTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGCGCGTTGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
DB 301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGGTAGAACGAGCGGTGCAAGCGCTGTAAGCGGCGGTGCACAAATC 420
DB 361 TGGTGGTGTGATGGTAGAACGAGCGGTGCAAGCGCTGTAAGCGGCGGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCGCTGGATGACACGAGTGCCTCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCGCTGGATGACACGAGTGCCTCA 480
QY 481 TTGCTGTGGAGCTGCTGCACATAATGTTCCGCGGTATTTCTTGATGTCTCTGACACAGA 540
DB 481 TTGCTGTGGAGCTGCTGCACATAATGTTCCGCGGTATTTCTTGATGTCTCTGACACAGA 540
QY 541 CACCCATCAACAGTATATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGTTCGATTTGGGTACACGAGCAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
DB 601 TGTTCGATTTGGGTACACGAGCAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
QY 661 CGCGTCTGCTCTGGCTGGCTGCGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
DB 661 CGCGTCTGCTCTGGCTGGCTGCGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAAACGGAAGCGGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAAACGGAAGCGGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAAACGATCAGATGGCGTGGCGCAAA 840
DB 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAAACGATCAGATGGCGTGGCGCAAA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
DB 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTCTCTGGGGCAAAACAGCGGTGGACCGTGTGTCGAATCTCTCAGGGCCAGCGGTGA 1020
DB 961 GCCTCTCTGGGGCAAAACAGCGGTGGACCGTGTGTCGAATCTCTCAGGGCCAGCGGTGA 1020

```

Qy	1021	AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAAACACCCCTGGCGGCCCAATA	1080
Db	1021		
Qy	1081	CGAAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGACAGAGTTT	1140
Db	1081		
Qy	1141	CCCGACTGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
Db	1141		
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATGCACTGCAACGGTGACCAATGCTTCTGGCG	1260
Db	1201		
Qy	1261	TCAGGACGCATCGGAAGCTGTGGTATGCGTGTGCGAGTCTGTAATCACTGCATAATTTCG	1320
Db	1261		
Qy	1321	TGTCGCTCAAGCGGCACATCCCGTCTCGGATTAATGTTTTTGGCGGACATCATACGGTT	1380
Db	1321		
Qy	1381	CTGGCAAAATATCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTTGGA	1440
Db	1381		
Qy	1441	ATTGTGCGGATTAACAAATTTACACAGGAACAGCCAGTCCGTTTAGTGTGTTTACAGA	1500
Db	1441		
Qy	1501	GCATTTCACCAACAGGACCATAG-----	1524
Db	1501		
Qy	1525	-----ATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG	1560
Db	1525		
Qy	1561	ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620
Db	1561		
Qy	1621	ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1680
Db	1621		
Qy	1681	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGCCAGGTT	1740
Db	1681		
Qy	1741	GCGCAACTGGCGATGGCCCTGACATTAATCTTGGGCACACGACCGCTTTGGTGGCTAC	1800
Db	1741		
Qy	1801	GCTCAATCTGGCGTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAGCTGTAT	1860
Db	1801		
Qy	1861	CGTTTACCTGGGATGCGGTACGTTTACACGCGAAGCTGATTCTTACCGGATCGCTGTT	1920
Db	1861		
Qy	1921	GAAGCGTTATCGCTGATTATATAAACAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1980
Db	1921		
Qy	1981	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	2040
Db	1981		
Qy	2041	CTGCAAGAACCGTACTTCACCTGGCGGCTGATTGCTGACGGGGGTTATGCGTTCAAG	2100
Db	2041		

Search completed: August 1, 2005, 17:22:13
Job time : 9284 secs

Qy	2041	TATGAAAAACGGCAGTACGACATTAAGACGTGGCGTGGATACGCTGGCGCGAAGCG	2100
Db	2101		
Qy	2101	TATGAAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATACGCTGGCGCGAAGCG	2160
Db	2101		
Qy	2161	GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAACACATGAATGACAGACACCGATTAC	2220
Db	2161		
Qy	2221	TCCATCGCAGAAAGCTGCTCTTTAATAAGGGGAAAAACAGGATGACATCAACAGGCCCGTGG	2280
Db	2221		
Qy	2281	GCATGGTCCACATCGACACACAGCAAGTGAATTAATGGTGTAAACGGTACTCCGACCTTC	2340
Db	2281		
Qy	2341	ANGGGTCAACATCGACACAGCAAGTGAATTAATGGTGTAAACGGTACTCCGACCTTC	2400
Db	2341		
Qy	2401	CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGTGACTGATGAAGGTCG	2460
Db	2401		
Qy	2461	GAAGCGGTTAATAAGACAAACCCGCTGGGTCCGCTAGCGTGAAGTCTTACGAGGAAG	2520
Db	2461		
Qy	2521	TTGGCGAAAGATCCACGTTATTCGCCACCATGGAACACGCCAGAAAGGTGAATCATG	2580
Db	2521		
Qy	2581	CCGAAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2640
Db	2581		
Qy	2641	GCCAGCGGTCGTACAGCTGTTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2700
Db	2641		
Qy	2701	GAATCCTCT 2709	2760
Db	2761	GGATCCTCT 2769	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 14:24:04 ; Search time 1094.6 Seconds
(without alignment)
17846.836 Million cell updates/sec

Title: US-09-765-555B-18
Perfect score: 3300
Sequence: 1 ccgacaccatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	4 AAD11592	Aad11592 Partial s
2	3232.8	98.0	3300	4 AAD11591	Aad11591 Partial s
3	3226.4	97.8	3300	4 AAD11588	Aad11588 Partial s
4	3226.4	97.8	3300	4 AAD11589	Aad11589 Partial s
5	3216.8	97.5	3300	4 AAD11590	Aad11590 Partial s
6	2705.8	82.0	8101	4 AAH25681	AAH25681 Nucleotid
7	2705.8	82.0	8101	6 ABL53238	ABL53238 Nucleotid
8	2690.4	81.5	7259	12 ADO23608	ADO23608 DNA encod
9	2690.4	81.5	7322	12 ADO23613	ADO23613 DNA encod
10	2690.4	81.5	7352	12 ADO23603	ADO23603 DNA encod
11	2690.4	81.5	7370	12 ADO23598	ADO23598 DNA encod
12	2690.4	81.5	7370	12 ADO23639	ADO23639 DNA encod
13	2690.4	81.5	7370	12 ADO23649	ADO23649 DNA encod
14	2690.4	81.5	7370	12 ADO23644	ADO23644 DNA encod
15	2690.4	81.5	7403	12 ADO23594	ADO23594 DNA encod
16	2690.4	81.5	7442	12 ADO23590	ADO23590 DNA encod
17	2690.4	81.5	7478	12 ADO23584	ADO23584 DNA encod
18	2688.8	81.5	7112	12 ADO23618	ADO23618 DNA encod
19	2636.8	79.9	6806	6 ABL49925	ABL49925 Maltose b
20	2605.4	79.0	7553	12 ADO23588	ADO23588 DNA encod

21	1609.4	48.8	5558	12 ADL90419	Adl90419 Clostridi
22	1474	44.7	4700	9 ACF06053	Act06053 Vector pJ
23	1472.4	44.6	4700	12 ADL18670	Adl18670 Vector pJ
24	1462.4	44.3	4920	10 ADH73598	Adh73598 Plasmid p
25	1462.4	44.3	4935	10 ADH73598	Adh73598 Plasmid p
26	1462.4	44.3	4945	10 ADH73600	Adh73600 Plasmid p
27	1462.4	44.3	4951	10 ADH73601	Adh73601 Plasmid p
28	1418.2	43.0	5903	8 ABZ23939	Abz23939 Nucleotid
29	1417.6	43.0	1511	6 ABQ73200	Abq73200 E. coli t
c 30	1417.6	43.0	5926	2 AAV32977	Aav32977 Tn7 donor
c 31	1417.6	43.0	5926	6 AAD45059	Aad45059 Transposo
c 32	1417.6	43.0	5926	12 ADG46817	Adg46817 Donor pla
33	1362	41.3	5201	12 ADL72229	Adl72229 DNA seque
34	1362	41.3	5201	12 ADL72228	Adl72228 DNA seque
c 35	1238.4	37.5	5024	9 ACF06051	Act06051 Plasmid p
c 36	1238.4	37.5	5024	12 ADL18668	Adl18668 Plasmid p
37	1237.2	37.5	1922	6 ABQ73089	Abq73089 MBP and z
38	1237.2	37.5	1922	10 ADF83622	Adf83622 MBP-human
39	1204	36.5	4557	2 AAT90491	Aat90491 Vector pl
40	1204	36.5	4969	12 ADO33522	Ado33522 Expressio
41	1204	36.5	5006	10 ADE24112	Ade24112 plasmid p
42	1204	36.5	5064	12 ADO33532	Ado33532 Vector pG
43	1204	36.5	5082	12 ADO33531	Ado33531 Vector pG
44	1204	36.5	6259	6 AAD29720	Aad29720 Plasmid p
45	1204	36.5	6823	3 AAC55459	Aac55459 Destinati

ALIGNMENTS

RESULT 1

AAD11592

ID AAD11592 standard; DNA; 3300 BP.

XX AAD11592;

XX 24-SEP-2001 (first entry)

XX Partial sequence of pMal-Ap3 and ZFPAP3 DNA.

XX Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;

XX modulation; plant technology; agriculture; Ap3; APETALA3; ds.

XX Unidentified.

XX Key Location/Qualifiers

FT CDS 2719..3270

FT /*tag= a

FT /product= "ZFPAP3 protein"

FT /transl_except= {pos:3124..3126, aa:Ser}

FT /transl_except= {pos:3127..3129, aa:Ser}

FT /transl_except= {pos:3136..3138, aa:Val}

FT /note= "CDS does not include start and stop codon"

FT /partial

FT primer_bind 2740..2790

FT /*tag= b

FT /bound_moiety= "F1-f2 primer"

FT primer_bind 2770..2850

FT /*tag= c

FT /bound_moiety= "F1-f1 primer"

FT primer_bind 2824..2889

FT /*tag= d

FT /bound_moiety= "F2-b primer"

FT primer_bind 2867..2940

FT /*tag= e

FT /bound_moiety= "F2-f primer"

FT primer_bind 2916..2973

FT /*tag= f

FT /bound_moiety= "F3-b1 primer"

FT primer_bind 2953..3021

FT /*tag= g

FT /bound_moiety= "F3-b2 primer"

FT primer_bind 2992..3042

```

FT      /*tag= h
FT      /bound_moiety= "F4-f2 primer"
FT      3022..3102
FT      /*tag= i
FT      /bound_moiety= "F4-f1 primer"
FT      3076..3141
FT      /*tag= j
FT      /bound_moiety= "F5-b primer"
FT      3119..3192
FT      /*tag= k
FT      /bound_moiety= "F5-f primer"
FT      3168..3225
FT      /*tag= l
FT      /bound_moiety= "F6-b1 primer"
FT      3205..3273
FT      /*tag= m
FT      /bound_moiety= "F6-b2 primer"
XX
PN      WO200152620-A2.
PD      26-JUL-2001.
XX
XX      19-JAN-2001; 2001WO-US001817.
XX
XX      21-JAN-2000; 2000US-0177468P.
XX      21-JUL-2000; 2000US-00620897.
XX
XX      (SCRI ) SCRIPPS RES INST.
XX      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI      Barbas CF, Stege JT, Guan X, Dalmia B;
XX
XX      WPI; 2001-465325/50.
XX      P-PSDB; AA06005.
XX
XX      New zinc finger proteins, useful for modulating or regulating gene
XX      expression and metabolic pathways in plants, e.g. for treating in the
XX      plant cells a disorder that is associated with abnormal expression of the
XX      target gene.
XX
XX      Example 4; Page 148-149; 156pp; English.
XX
XX      The patent discloses methods and compositions to modulate the expression
XX      of a target gene in plant cells. The method involves providing plant
XX      cells with a zinc finger protein (ZFP) which is capable of specifically
XX      binding to a target nucleotide sequence or its complementary strand
XX      within a target gene and allowing the ZFP binding to the target
XX      nucleotide sequence, where the expression of the target gene in the plant
XX      cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX      for modulating or regulating gene expression and metabolic pathways in
XX      plants. The ZFP, fusion proteins and methods are useful in plant and
XX      agricultural technology. The method is useful particularly for treating a
XX      disorder in the plant cells, where the disorder is associated with
XX      abnormal expression of the target gene. The present DNA sequence is the
XX      partial sequence of pMal-AP3 (AP3FLA3) and ZFPp3 DNA
XX
SQ      Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;
Query Match
Best Local Similarity 100.0%; Score 3300; DB 4; Length 3300;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CCGACACCATCGAATGGTGAACACCTTTCGCGTATGGCATGATAGCCCGCGGAAGAGA 60
Db      1 CCGACACCATCGAATGGTGAACACCTTTCGCGTATGGCATGATAGCCCGCGGAAGAGA 60
Qy      61 GTCAATTCAGGTGGTGAATGTGAACACAGTATACGTTATACGATGTCGAGATATCCG 120
Db      61 GTCAATTCAGGTGGTGAATGTGAACACAGTATACGTTATACGATGTCGAGATATCCG 120
Qy      121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGTTCCTGCCGAAA 180
Db      121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGTTCCTGCCGAAA 180

```

```

Qy      181 CGCGGGAAGAGTCGAAGCGCGGATGGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Db      181 CGCGGGAAGAGTCGAAGCGCGGATGGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Qy      241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTTGC 300
Db      241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTTGC 300
Qy      301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db      301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Qy      361 TGGTGGTGTGATAGAACGAGCGCGCTGGAAGCGCTGTAAGCGCGCGGTGCACAATC 420
Db      361 TGGTGGTGTGATAGAACGAGCGCGCTGGAAGCGCTGTAAGCGCGCGGTGCACAATC 420
Qy      421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACCTATCCGCTGATGACCAAGATGCCA 480
Db      421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACCTATCCGCTGATGACCAAGATGCCA 480
Qy      481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGGTTAATTTCTTGATGTCTCTGACCAGA 540
Db      481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGGTTAATTTCTTGATGTCTCTGACCAGA 540
Qy      541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTAGCGACTGGGCGTGGAGCATC 600
Db      541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTAGCGACTGGGCGTGGAGCATC 600
Qy      601 TGGTCGCATTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db      601 TGGTCGCATTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Qy      661 CGGTCCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db      661 CGGTCCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy      721 CGGAACGGGAAGCGAGCTGGAGTGCCATGTCGGTTCCTCAACCAACCATCAAAATGCTGA 780
Db      721 CGGAACGGGAAGCGAGCTGGAGTGCCATGTCGGTTCCTCAACCAACCATCAAAATGCTGA 780
Qy      781 ATGAGGCAATCGTTCCCACTCGCATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db      781 ATGAGGCAATCGTTCCCACTCGCATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Qy      841 TGGCGCCATTACCGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db      841 TGGCGCCATTACCGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy      901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCACCATCAAAACAGATTTTC 960
Db      901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCACCATCAAAACAGATTTTC 960
Qy      961 GCCTGCTGGGGCAAAACAGGTGGACCGTGTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db      961 GCCTGCTGGGGCAAAACAGGTGGACCGTGTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy      1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db      1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Qy      1081 CGCAAAACCGCTCTCCCGCGCGTTGGCGGATTCATTAAATGCAGCTGGGACGACAGGTTT 1140
Db      1081 CGCAAAACCGCTCTCCCGCGCGTTGGCGGATTCATTAAATGCAGCTGGGACGACAGGTTT 1140
Qy      1141 CCCGACTGGAACCGGGGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Db      1141 CCCGACTGGAACCGGGGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Qy      1201 GCACAATTTCTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db      1201 GCACAATTTCTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260

```

Qy	1261	TCAGGCAGCCATTCGGAAAGCTGTGTGTATGGCTGTGCAGGTCGTAAATCACTGCAATAATTCG	1320
Db	1261	TCAGGCAGCCATTCGGAAAGCTGTGTGTATGGCTGTGCAGGTCGTAAATCACTGCAATAATTCG	1320
Qy	1321	TGTCGCTCAAGGGCGACCTCCCGTTCTCGGATAATGTTTTTTTTCGCGCCGACATCAATAACGGTT	1380
Db	1321	TGTCGCTCAAGGGCGACCTCCCGTTCTCGGATAATGTTTTTTTTCGCGCCGACATCAATAACGGTT	1380
Qy	1381	CTGGCAAAATATCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
Db	1381	CTGGCAAAATATCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
Qy	1441	ATTGTGAGCGGATTAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTTTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATTAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTTTTTTCACGA	1500
Qy	1501	GCATTCACCAACAAAGGACCATAGATTATGAAACTGAAGAAGTAACCTGTAATCTCG	1560
Db	1501	GCATTCACCAACAAAGGACCATAGATTATGAAACTGAAGAAGTAACCTGTAATCTCG	1560
Qy	1561	ATTAACGGCGATTAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAACGGCGATTAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGAAATTAAGGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACCAAGTT	1680
Db	1621	ACCGAAATTAAGGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACCAAGTT	1680
Qy	1681	GGCGCACTGGCGATGGCCCTGCATTTATCTTCTGGGCAACAGACCGCTTTGGTGCTAC	1740
Db	1681	GGCGCACTGGCGATGGCCCTGCATTTATCTTCTGGGCAACAGACCGCTTTGGTGCTAC	1740
Qy	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAAAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCCGTAACGTTTAAACGGCAAGCTGATACCGGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCCGTAACGTTTAAACGGCAAGCTGATACCGGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTAACAAGATCTGCTGCCGAACCGCCCAAAACCTGGGA	1920
Db	1861	GAAGCGTTATCGCTGATTAACAAGATCTGCTGCCGAACCGCCCAAAACCTGGGA	1920
Qy	1921	GAGATCCCGCGCTGGATAAGAACTCAAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAGAACTCAAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACGGGGTATTCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACGGGGTATTCGTTCAAG	2040
Qy	2041	TATGAAAAACCGCAAGTACGACATTTAAAGACGTGGGGCTGATTAACCGCTGGCGGAAGCG	2100
Db	2041	TATGAAAAACCGCAAGTACGACATTTAAAGACGTGGGGCTGATTAACCGCTGGCGGAAGCG	2100
Qy	2101	GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGATGCAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGATGCAGACACCGATTAC	2160
Qy	2161	TCCATCGCAGAGCTGCTCTTTTAAATAAGGGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAGAGCTGCTCTTTTAAATAAGGGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Qy	2221	GCATGGTCCAAACATTCGACACACAGCAAAAGTGAAATTAATGGTGTAACGGTACTTCGCACTTC	2280
Db	2221	GCATGGTCCAAACATTCGACACACAGCAAAAGTGAAATTAATGGTGTAACGGTACTTCGCACTTC	2280
Qy	2281	AAGGGTCAACCATCCAAACCGTTCGTTGGCGTCTGAGCCGAGGTATTAACGGCCGACGT	2340
Db	2281	AAGGGTCAACCATCCAAACCGTTCGTTGGCGTCTGAGCCGAGGTATTAACGGCCGACGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTTATCTGTTGACTGATGAAGGTCGT	2400

RESULT 2

AAD11591

AAD11591
ID AAD11591 standard: DNA: 3300 BP.

ID AAD11591;
XX
AC AAD11591;

DT 24-SEP-2001 (first entry)

XX
DE partial sequence of pMal-m4 and ZFPm4 DNA.

XX Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.
XX Unidentified.

Key Location/Qualifiers

FT CDS 2719..3270
FT /tag= a
FT /product= "ZFPm4 protein"
FT /transl_except= (pos:3046..3048, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /tag= m
FT /bound_moiety= "F6-b2 primer"

XX WO200152620-A2.

XX PD 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US001817.

XX 21-JAN-2000; 2000US-0177458P.

XX 21-JUL-2000; 2000US-00620897.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA AGRIC DISCOVERY INC.

XX Barbas CF, Stege JT, Guan X, Dalmia B;

XX WPI; 2001-465325/50.

XX P-PSDB; AAE06004.

XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.

XX Example 4; Page 145-147; 156pp; English.

XX The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m4 and ZFPm4 DNA

XX SQ Sequence 3300 BP; 849 A; 865 C; 880 G; 706 T; 0 U; 0 Other;

Query Match 98.0%; Score 3232.8; DB 4; Length 3300;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 3258; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db |||
Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db |||
Qy 61 GTCAATTACGGGTGGTGAATGTGAACACCACTAAGTTATACGATGTCGAGAGTATGCCG 120
Db |||
Qy 61 GTCAATTACGGGTGGTGAATGTGAACACCACTAAGTTATACGATGTCGAGAGTATGCCG 120
Db |||
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCCAGCCAGCTTCTTCGCAAAA 180
Db |||
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCCAGCCAGCTTCTTCGCAAAA 180
Db |||
Qy 181 CGCGGGAAAAGTGGAGCGCGATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db |||
Qy 181 CGCGGGAAAAGTGGAGCGCGATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db |||
Qy 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGGCGTTCGCCACTCCAGTCTGGCCCTGC 300
Db |||
Qy 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGGCGTTCGCCACTCCAGTCTGGCCCTGC 300
Db |||
Qy 301 ACGCCGCTCGCAAAATGTCGCGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db |||
Qy 301 ACGCCGCTCGCAAAATGTCGCGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db |||
Qy 361 TGGTGTGTGTCGATGAGTGAACGAGCGGTGCGAAGCGCTGTAAGCGCGGTGCACAATC 420
Db |||
Qy 361 TGGTGTGTGTCGATGAGTGAACGAGCGGTGCGAAGCGCTGTAAGCGCGGTGCACAATC 420
Db |||
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
Db |||
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
Db |||
Qy 481 TTGCTGTGGAGCTGCTGCATTAATGTTCCGCGGTATTTCTTGATGTCTCTGACAGA 540
Db |||
Qy 481 TTGCTGTGGAGCTGCTGCATTAATGTTCCGCGGTATTTCTTGATGTCTCTGACAGA 540
Db |||
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGAGCGGTACGCGCTGGGCGTGGAGCATC 600
Db |||
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGAGCGGTACGCGCTGGGCGTGGAGCATC 600
Db |||
Qy 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db |||
Qy 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db |||
Qy 661 CGCGTCTGGCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
Db |||
Qy 661 CGCGTCTGGCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
Db |||
Qy 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db |||
Qy 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db |||

QY 781 ATGAGGCGATCGTTCCCACTCGATGCTGTTGCCAAGCATCAGATGCGCTGGGCGCAA 840
DB 781 ATGAGGCGATCGTTCCCACTCGATGCTGTTGCCAAGCATCAGATGCGCTGGGCGCAA 840
QY 841 TGGCGCCATTACCGAGTCCGGGCTCGCGTGTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB 841 TGGCGCCATTACCGAGTCCGGGCTCGCGTGTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCGAAGACAGCTGAGCCGCTTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCGAAGACAGCTGAGCCGCTTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGGTGGCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTGGTGGCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAATCCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGCTGGCGACGACAGGTTT 1140
DB 1081 CGCAATCCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGCTGGCGACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGCGGTCACCAATGCTTCGGCG 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGCGGTCACCAATGCTTCGGCG 1260
QY 1261 TCAGGCGAGCATCGGAAGCTGGTATGGCTGTGCGAGTTCGTAATCACTGCGATTAATTCG 1320
DB 1261 TCAGGCGAGCATCGGAAGCTGGTATGGCTGTGCGAGTTCGTAATCACTGCGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCGGATTAATGTTTTGCGCGGACATCATACAGGTT 1380
DB 1321 TGTGCTCAAGCGCACTCCCGTTCGGATTAATGTTTTGCGCGGACATCATACAGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGTGA 1440
DB 1381 CTGGCAATATTTCTGAATAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGTGA 1440
QY 1441 ATTGTAGCGGATACAATTTACACAGGAAACAGCCAGTCCGTTAGGTTTTCACGA 1500
DB 1441 ATTGTAGCGGATACAATTTACACAGGAAACAGCCAGTCCGTTAGGTTTTCACGA 1500
QY 1501 GCACCTCACCAACAGGACCATAGATTATGAACCTGAAGAGTTAACTGGTAACTCTGG 1560
DB 1501 GCACCTCACCAACAGGACCATAGATTATGAACCTGAAGAGTTAACTGGTAACTCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCGCCAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCGCCAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTCTGGGCACAGCCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTCTCTGGGCACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCCGCCAAAACCTGGGAA 1920

DB 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGTCTGACGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGTCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
DB 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAAAAACAACATGAAATGCAAGACCGGATTAC 2160
DB 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAAAAACAACATGAAATGCAAGACCGGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAAGCGAACAACAGCGATGACCATCAACGGCCCGTGG 2220
DB 2161 TCCATCGCAGAGCTGCGCTTTAATAAAGCGAACAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAACTCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCGACCTTC 2280
DB 2221 GCATGGTCCAACTCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCGACCTTC 2280
QY 2281 AAGGGTCAACCATCAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACCGCCCACTG 2340
DB 2281 AAGGGTCAACCATCAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACCGCCCACTG 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
DB 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTCTTACGAGGAAG 2460
DB 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTCTTACGAGGAAG 2460
QY 2461 TTGCGAAGAGTCCAACGTTATTCGCGCCCAACCATGGAACAAACGCGCAGAAAGGTGAATCATG 2520
DB 2461 TTGCGAAGAGTCCAACGTTATTCGCGCCCAACCATGGAACAAACGCGCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACAATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCGTATGCGGTGATCAACGCC 2580
DB 2521 CCGAACAATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCGTATGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
DB 2641 AACCAACAACAATAAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
QY 2701 GGATCCTCTTCTCTGTTGGCCAGCGGCCCTCGAGCCCGGGGAGAACCCCTATGCTGT 2760
DB 2701 GGATCCTCTTCTCTGTTGGCCAGCGGCCCTCGAGCCCGGGGAGAACCCCTATGCTGT 2760
QY 2761 CCGGAATGTTGTTAGTCTCTCAGCCAGAGAGCTCCCTGTTGGCGCCACAGCGTACCAC 2820
DB 2761 CCGGAATGTTGTTAGTCTCTCAGCCAGAGAGCTCCCTGTTGGCGCCACAGCGTACCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTCCGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
DB 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTCCGCAAAATCTTTTAGCCAGAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGTCGAATGTGGC 2940
DB 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGTCGAATGTGGC 2940
QY 2941 AAGTCTTTACGCCAGTCCAGCAACCTGTTGGCGCCCAACCGTACTCACACCGGGAGAG 3000

```

Db 2941 AAGTCTTCAGTGATTGTCGTGATCTTCGAGGACCAACGTACTCACACCGGGGAGAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTTGTAAGTCCTTCAGCACGAGTGGCTCCTTGTTAGACAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTTGTAAGTCCTTCAGAGCTCTCACCTGTGGCCAC 3060
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTACG 3120
Db 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTACG 3120
QY 3121 CAGCGGGCCCACTCGAAGCCCATCAAGCACTCATCTACTGGGAGAACCATACAAATGT 3180
Db 3121 CGACGGAATACTGGTGGCCATCAAGCACTCATCTACTGGGAGAACCATACAAATGT 3180
QY 3181 CCAGATGTGGCAAGTCTTCTCAACTTCAGGCAACTTGGTCCGTCACCAAGTACTCAC 3240
Db 3181 CCAGATGTGGCAAGTCTTCTCAACTTCAGGCAACTTGGTCCGTCACCAAGTACTCAC 3240
QY 3241 ACCGTTAAAAAACTAGTGGCGAGCGCCGCGGCGAGTACCGTACGAGCTTCGGACTACGCT 3300
Db 3241 ACCGTTAAAAAACTAGTGGCGAGCGCCGCGGCGAGTACCGTACGAGCTTCGGACTACGCT 3300

RESULT 3
AAD11588
ID AAD11588 standard; DNA; 3300 BP.
AC
XX
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m1 and ZFPm1 DNA.
XX
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key
FT Location/Qualifiers
FT 2719..3270
FT /tag= a
FT /product= "ZFPm1 protein"
FT /transl_except= (pos:2785..2787, aa:Asp)
FT /transl_except= (pos:2788..2790, aa:Pro)
FT /transl_except= (pos:2791..2793, aa:Gly)
FT /transl_except= (pos:2872..2874, aa:Arg)
FT /transl_except= (pos:2875..2877, aa:Ala)
FT /transl_except= (pos:2878..2880, aa:His)
FT /transl_except= (pos:2884..2886, aa:Glu)
FT /transl_except= (pos:2953..2955, aa:Gln)
FT /transl_except= (pos:2959..2961, aa:Ser)
FT /transl_except= (pos:3121..3123, aa:Arg)
FT /transl_except= (pos:3124..3126, aa:Ser)
FT /transl_except= (pos:3127..3129, aa:Asp)
FT /transl_except= (pos:3130..3132, aa:Asn)
FT /transl_except= (pos:3136..3138, aa:Val)
FT /transl_except= (pos:3139..3141, aa:Arg)
FT /transl_except= (pos:3205..3207, aa:Gln)
FT /transl_except= (pos:3208..3210, aa:Ala)
FT /transl_except= (pos:3211..3213, aa:Gly)
FT /transl_except= (pos:3214..3216, aa:His)
FT /transl_except= (pos:3220..3222, aa:Ala)
FT /transl_except= (pos:3223..3225, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /tag= d

```

```

FT primer_bind
FT /bound_moiety= "F2-b primer"
FT 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"
FT
XX WO200152620-A2.
PN
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US001817.
XX
XX 21-JAN-2000; 2000US-0177468P.
PR
XX 21-JUL-2000; 2000US-00620897.
XX
XX (SCRI ) SCRIPPS RES INST.
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CF, Stege JT, Guan X, Dalmia B;
XX WPI; 2001-465325/50.
XX P-PSDB; AAE06000.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX
XX Example 4; Page 138-139; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX partial sequence of pMal-m1 and ZFPm1 DNA
XX
XX Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;

```

```

Query Match          97.8%; Score 3226.4; DB 4; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGCGCATGTAGCGCCGGAAGAGA 60
Db |||||
1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGCGCATGTAGCGCCGGAAGAGA 60
QY 61 GTCAATTTCAGGGTGTGTAATGTGA AACACAGTAACGTTATACGATGTCCGACAGTATGCCG 120
Db |||||
61 GTCAATTTCAGGGTGTGTAATGTGA AACACAGTAACGTTATACGATGTCCGACAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGA ACCAGGCCAGCCACGTTTCTGCGAAAA 180
Db |||||
121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGA ACCAGGCCAGCCACGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGAGCGGAGCTGAATTACATTCGCCAACCGCTGGGCAC 240
Db |||||
181 CGCGGAAAAAGTGAAGCGCGATGAGCGGAGCTGAATTACATTCGCCAACCGCTGGGCAC 240
QY 241 AACAACTGCGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db |||||
241 AACAACTGCGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGGCGCTCGCAAAATTGTCGGCGGATTAATCTCGCGCGATCAA CTGGGTGCCAGC 360
Db |||||
301 ACGGCGCTCGCAAAATTGTCGGCGGATTAATCTCGCGCGATCAA CTGGGTGCCAGC 360
QY 361 TGGTGGTGTGATGGTGA AACGAGCGCGTGAAGCCGTTAAAGCGCGGTGCAACAATC 420
Db |||||
361 TGGTGGTGTGATGGTGA AACGAGCGCGTGAAGCCGTTAAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCGCTGATGACACAGGATGCCA 480
Db |||||
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCGCTGATGACACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA 540
Db |||||
481 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA 540
QY 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db |||||
541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGCAATGGGTCAACAGCAAAATCGCGTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db |||||
601 TGGTCGCAATGGGTCAACAGCAAAATCGCGTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTCGCTGGCTGGCATAAATCTCACTCGCAATCAAATTCAGCCGATAG 720
Db |||||
661 CGCGTCTCGCTCGCTGGCTGGCATAAATCTCACTCGCAATCAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db |||||
721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGATGCTGTTGGTTCGCAACGATCAGATGCGCTGGCGCGAA 840
Db |||||
781 ATGAGGGCATGTTCCCACTCGATGCTGTTGGTTCGCAACGATCAGATGCGCTGGCGCGAA 840
QY 841 TGC CGGCCATTAACCGAGTCCGGGCTCGCGTGTGGTTCGGATATCTCGGTAGTGGGATAG 900
Db |||||
841 TGC CGGCCATTAACCGAGTCCGGGCTCGCGTGTGGTTCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACGATCATGTTATATCCCGCGTTAACCCACATCAAAACAGGATTTTC 960
Db |||||
901 ACGATACCGAAGACGATCATGTTATATCCCGCGTTAACCCACATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db |||||
961 GCCTGCTGGGCAACACGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGA AAAAAGAAACCAACCCCTGGCGCCCAATA 1080
Db |||||
1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGA AAAAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCCTCTCCCGCGCGTTGGCGGATTCATTAATGCACTGGCA CGACAGGTTT 1140
Db |||||

Db 1081 CGCAAAACCGCCTCTCCCGCGCGTTGGCGGATTCATTAATGCACTGGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db |||||
1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATGCACTGCACGCGTGCACCAATGCTTCTGGC 1260
Db |||||
1201 GCACAAATCTCATGTTTGACAGCTTATCATGCACTGCACGCGTGCACCAATGCTTCTGGC 1260
QY 1261 TCAGCGAGCCATCCGAAAGCTGTGATGCGTGTGCGAGTCCGTAATCACTGCAATAATTCG 1320
Db |||||
1261 TCAGCGAGCCATCCGAAAGCTGTGATGCGTGTGCGAGTCCGTAATCACTGCAATAATTCG 1320
QY 1321 TGTCTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGTT 1380
Db |||||
1321 TGTCTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db |||||
1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACCAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Db |||||
1441 ATTGTGAGCGGATAACCAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCA TAGATTTATGAAAACTGAAGAGGTAAATCTGTAATCTGG 1560
Db |||||
1501 GCACCTTCAACCAAGGACCA TAGATTTATGAAAACTGAAGAGGTAAATCTGTAATCTGG 1560
QY 1561 ATTTAAACCGCGATTAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db |||||
1561 ATTTAAACCGCGATTAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
Db |||||
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
QY 1681 CGCGCAACTGCGCATGCGCTTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db |||||
1681 CGCGCAACTGCGCATGCGCTTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGAGCAAGCTGTAT 1800
Db |||||
1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGAGCAAGCTGTAT 1800
QY 1801 CCGTTTACCTCGGATGCGTACGTTACAAACGCAAGCTGATTTGCTTACC CGATCGCTGTT 1860
Db |||||
1801 CCGTTTACCTCGGATGCGTACGTTACAAACGCAAGCTGATTTGCTTACC CGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAAACCTGGGAA 1920
Db |||||
1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db |||||
1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGCGGGTTATGCGTTCAAG 2040
Db |||||
1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGCGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db |||||
2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGGTGGCTGATTTAAAAAACAACATGAATGCAAGACCGGATTAC 2160
Db |||||
2101 GGTCTGACCTTCTCGTGGTGGCTGATTTAAAAAACAACATGAATGCAAGACCGGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCTCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCCGTTG 2220
Db |||||

Db	2161	TCCATCGCAGAGCTGCCTTTAATAAGGGCGAAACAGCGATGACCATCAACGGCCCGGTG	222
Qy	2221	GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCACCTTC	2280
Db	2221	GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCACCTTC	2280
Qy	2281	AAGGGTCAACCATCGAAACCGTTTCGTTGGCGTGTGAGCGCAGGTATTTAAACGCCCGCAGT	2340
Db	2281	AAGGGTCAACCATCGAAACCGTTTCGTTGGCGTGTGAGCGCAGGTATTTAAACGCCCGCAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACCTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACCTATCTGCTGACTGATGAAGTCTG	2400
Qy	2401	GAAGCGGTTAATAAGAGCAAAACCGCTGGGTGCCGTAGCGCTGAAAGTCTTTACGAGGAAG	2460
Db	2401	GAAGCGGTTAATAAGAGCAAAACCGCTGGGTGCCGTAGCGCTGAAAGTCTTTACGAGGAAG	2460
Qy	2461	TTGGCGGAAAGATCCACGTATTTGCGGCCACCAATGAGAAACGCCCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGGAAAGATCCACGTATTTGCGGCCACCAATGAGAAACGCCCCAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCCGTGCGTACTCGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCCGTGCGTACTCGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG	2640
Qy	2641	AACAACCAACAAATAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2641	AACAACCAACAAATAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Qy	2701	GGATCCTCTTCCTCTGTGGCCCCAGCGGCCCTCGAGCCGGGGAGAGCCCTATGCTTGT	2760
Db	2701	GGATCCTCTTCCTCTGTGGCCCCAGCGGCCCTCGAGCCGGGGAGAGCCCTATGCTTGT	2760
Qy	2761	CCGGAATGTGGTAACTCCTTCAGCGCAGAGCAGCTCCCTGGTGGCCACACGCGTACCAC	2820
Db	2761	CCGGAATGTGGTAACTCCTTCAGAGCTCTCACCCTGGTGGCCACACGCGTACCAC	2820
Qy	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTCGGGCAAACTCTTTTACGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTCGGGCAAACTCTTTTACGCCAGTCCAGCAAC	2880
Qy	2881	CTGGTGGCCATCAACGCACTACTCTGGCGAGAGGCATACAAATGTCAGAAATGTGGC	2940
Db	2881	CTGGTGGCCATCAACGCACTACTCTGGCGAGAGGCATACAAATGTCAGAAATGTGGC	2940
Qy	2941	AAGTCTTTTCAGCCAGTCCAGCAACCTCGTGGCCACCAAGCTACTCAGCCGGGAGAG	3000
Db	2941	AAGTCTTTTCTCGTCTGCAATCTCGTCCGGCACCAGCTACTCAGCCGGGAGAG	3000
Qy	3001	CCCTATGCTTTGTCGGGAATGTGTAAGTCTCTTCAGCACCGAGTGGCTCTTTGGTTAGACAC	3060
Db	3001	CCCTATGCTTTGTCGGGAATGTGTAAGTCTCTTCAGCGCAGCGATAACTGCTGGTGGCCAC	3060
Qy	3061	CAGGTTACCCACACCGGTGAAAAACCGTATAAATGCCCAGAGTCGGCAAACTTTTTCAGC	3120
Db	3061	CAGGTTACCCACACCGGTGAAAAACCGTATAAATGCCCAGAGTCGGCAAACTTTTTCAGC	3120
Qy	3121	CAGCGCCCACTGGAAACCGCATCAACGCACTACTCGGCGAGAGCCATCAAAATGT	3180
Db	3121	CAGCGCGGCCACTGTCGCGCCATCAACGCACTACTCGGCGAGAGCCATCAAAATGT	3180
Qy	3181	CCAGAATGTGGCAAGTCTTTTCTCAACTTCAGGCAACTTGGTTCGGTCAACCAACGCTACTC	3240
Db	3181	CCAGAATGTGGCAAGTCTTTTCTCTCGTCTGCAAACTCTCGTCCGGCACCACGCTACTC	3240
Qy	3241	ACCGGTAACCACTAGTGGCGCAGCGCGCGCAGTACCCGTACGAGCTTCGGGCTACGCT	3300
Db	3241	ACCGGTAACCACTAGTGGCGCAGCGCGCGCAGTACCCGTACGAGCTTCGGGCTACGCT	3300

XX WPI, 2001-465325/50.
DR P-PSDB; AAE06002.
XX
PT New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 140-142; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m2 and ZFPm2 DNA
XX
SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
Query Match 97.8%; Score 3226.4; DB 4; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 1 CGGACACCATCAATGGTGGCAAAACCTTTCGGCGTATGGCATGATAGCCCGGAAGAGA 60
DB 1 CGGACACCATCAATGGTGGCAAAACCTTTCGGCGTATGGCATGATAGCCCGGAAGAGA 60
QY 61 GTCAATTCCAGGTGTGAATGTGAACAGTAACTGTTATACGATGTCGAGAGTATGCG 120
DB 61 GTCAATTCCAGGTGTGAATGTGAACAGTAACTGTTATACGATGTCGAGAGTATGCG 120
QY 121 GTGCTCTTATCAGACCCGTTTCCCGCGTGTGAACAGCCAGCCACGCTTTCGCGAAAA 180
DB 121 GTGCTCTTATCAGACCCGTTTCCCGCGTGTGAACAGCCAGCCACGCTTTCGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAAATTAACATTCGCCAACCCGCTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAAATTAACATTCGCCAACCCGCTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTGTGCTGATTTGGCGTTTGGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAACAGTGTGCTGATTTGGCGTTTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTGGCAAAATTTGGCGGCGATTAATACTCTCGCGCGATCAACTGGGTGGCAGCG 360
DB 301 ACGCGCGTGGCAAAATTTGGCGGCGATTAATACTCTCGCGCGATCAACTGGGTGGCAGCG 360
QY 361 TGGTGGTGTGATGTAGAACGAAAGCGCGTGAAGCCCTGTAAAGCGGCGTGCACAATC 420
DB 361 TGGTGGTGTGATGTAGAACGAAAGCGCGTGAAGCCCTGTAAAGCGGCGTGCACAATC 420
QY 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480
DB 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCTCACTAATGTTCGGCGGTATTCTTGATGCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCTCACTAATGTTCGGCGGTATTCTTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
QY 601 TGGTGGCATTTGGGTACCGAAATTCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTGGCATTTGGGTACCGAAATTCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG

DB 601 TGGTGGCATTTGGGTACCGAAATTCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCATCTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCATCTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
DB 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTTGCACACGATCAGATGCGCTGGCGGCA 840
DB 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTTGCACACGATCAGATGCGCTGGCGGCA 840
QY 841 TGGCGCCATTAACGAGTCCGGGCTGCGCGTGGTGGTGGATATCTCGTAGTGGATAGC 900
DB 841 TGGCGCCATTAACGAGTCCGGGCTGCGCGTGGTGGTGGATATCTCGTAGTGGATAGC 900
QY 901 ACGTACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC 960
DB 901 ACGTACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGTGTGCGCTCTCATCTGTTGAAAGAAAAACCACTTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGTGTGCGCTCTCATCTGTTGAAAGAAAAACCACTTGGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCTCCCGCGGTGGCGGATTCATTAATCGAGCTGGCAGCAGAGTTT 1140
DB 1081 CGCAACCGCTCTCTCCCGCGGTGGCGGATTCATTAATCGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTAGTCTACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTAGTCTACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGATGCGGTCGACCAATGCTTCTGCGC 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGATGCGGTCGACCAATGCTTCTGCGC 1260
QY 1261 TCAGGACCGCATCGGAAGCTGTGATGCTGTGCGAGTCTGTAATCATCTGATTAATTCG 1320
DB 1261 TCAGGACCGCATCGGAAGCTGTGATGCTGTGCGAGTCTGTAATCATCTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGCGACATCATACCGGTT 1380
DB 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGCGACATCATACCGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAAATGAGCTGTGTAACAATTAATCATCGGCTCGTATATGTGTGA 1440
DB 1381 CTGGCAAAATTTCTGAAAATGAGCTGTGTAACAATTAATCATCGGCTCGTATATGTGTGA 1440
QY 1441 ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGTAGTGTTCACGA 1500
DB 1441 ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGTAGTGTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAACCTGAAGAGGTAATACTGGTAACTCG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAACCTGAAGAGGTAATACTGGTAACTCG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACCGGCTCGCTGAAGTCCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACCGGCTCGCTGAAGTCCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740

```
QY 1741 GCTCAATCTGGCTGTGTGGCTGAATCAACCCGGAACAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAATCAACCCGGAACAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACTCGGATGCCGTAGCTTTACAAAGCGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACTCGGATGCCGTAGCTTTACAAAGCGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAAGCGTTATCGCTGATTTATAAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAAGCGTTATCGCTGATTTATAAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTCGGCGCTGATTGCTGTCAGCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTCGGCGCTGATTGCTGTCAGCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGCAAGACG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGCAAGACG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGCAACACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGCAACACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAATGGTGAACGGTACTCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAATGGTGAACGGTACTCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCCGAAACATCTGCTGACTGATGAAGGTCGTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCCGAAACATCTGCTGACTGATGAAGGTCGTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAAGTCTTACGAGGAAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAAGTCTTACGAGGAAAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTGCGCCACCATGGAACACGCCAGAACGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTGCGCCACCATGGAACACGCCAGAACGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTG 2640
Db 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGGGGAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGGGGAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGCGCGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGCGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTTTCAGCCAGCAGCTCCCTGCTGTCGCCACGAGCTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTTCAGCCAGCAGCTCCCTGCTGTCGCCACGAGCTACCCAC 2820
```

```
QY 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTCCGCGCAAACTCTTTTACGAGTCCAGCAAC 2880
Db 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTCCGCGCAAACTCTTTTACGAGTCCAGCAAC 2880
QY 2881 CTGCTGCGCCATCAACGCACTCATATCTGGGAGAGGCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGCTGCGCCATCAACGCACTCATATCTGGGAGAGGCATACAAATGTCCAGAAATGTGGC 2940
QY 2941 AAGTCTTTCTAGCCAGTCCAGCAACCTGCTGGGCCCAACAGTACTCAACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTAGCCAGTCCAGCAACCTGCTGGGCCCAACAGTACTCAACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTGGTAACTCTTCAAGCACCACTGCTCTTGGTTAGACAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAACTCTTCAAGCACCACTGCTCTTGGTTAGACAC 3060
QY 3061 CAGGTTACCCACACCGGGTGAACACCGTATAAATGCCAGAGTCCGCGCAAACTCTTTTACG 3120
Db 3061 CAGGTTACCCACACCGGGTGAACACCGTATAAATGCCAGAGTCCGCGCAAACTCTTTTACG 3120
QY 3121 CAGCGCCGCCACCTGGAAACCGCATCAACGCACTCATATCTGGCGAAGGCCATACAAATGT 3180
Db 3121 CAGCGCCGCCACCTGGCAACCGCATCAACGCACTCATATCTGGCGAAGGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTTCAGGCAAACTTGGTCCGTCAACCAACGTAAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTTCAGGCAAACTTGGTCCGTCAACCAACGTAAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCGCAGTACCCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCGCAGTACCCGTACGACGTTCCGGACTACGCT 3300

RESULT 5
AAD11590
ID AAD11590 standard; DNA; 3300 BP.
XX
AC AAD11590;
AC
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m3 and ZFPm3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key
CDS Location/Qualifiers
FT 2719..3270
FT /*tag= a
FT /product= "ZFPm3 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind
FT 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind
FT 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind
FT 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind
FT 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind
FT 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind
FT 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind
FT 2992..3042
```


FT primer_bind /*tag= h /bound_moiety= "F4-f2 primer" 3022..3102

FT primer_bind /*tag= i /bound_moiety= "F4-f1 primer" 3076..3141

FT primer_bind /*tag= j /bound_moiety= "F5-b primer" 3119..3192

FT primer_bind /*tag= k /bound_moiety= "F5-f primer" 3168..3225

FT primer_bind /*tag= l /bound_moiety= "F6-b1 primer" 3205..3273

FT primer_bind /*tag= m /bound_moiety= "F6-b2 primer"

XX WO200152620-A2.

PN 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US001817.

XX 21-JAN-2000; 2000US-0177469P.

PR 21-JUL-2000; 2000US-00620897.

XX (SCRI) SCRIPPS RES INST.

PA (SYCN) SYNGENTA AGRIC DISCOVERY INC.

XX Barbas CF, Stege JT, Guan X, Dalmia B;

PI WPI; 2001-465325/50.

DR P-PSDB; AA06003.

XX New zinc finger proteins, useful for modulating or regulating gene expression and metabolic pathways in plants, e.g. for treating in the plant cells a disorder that is associated with abnormal expression of the target gene.

XX Example 4; Page 143-145; 156pp; English.

XX The patent discloses methods and compositions to modulate the expression of a target gene in plant cells. The method involves providing plant cells with a zinc finger protein (ZFP) which is capable of specifically binding to a target nucleotide sequence or its complementary strand within a target gene and allowing the ZFP binding to the target nucleotide sequence, where the expression of the target gene in the plant cells is modulated. The ZFP and fusions of the ZFP proteins are useful for modulating or regulating gene expression and metabolic pathways in plants. The ZFP, fusion proteins and methods are useful in plant and agricultural technology. The method is useful particularly for treating a disorder in the plant cells, where the disorder is associated with abnormal expression of the target gene. The present DNA sequence is the partial sequence of pMal-m3 and ZFPm3 DNA

SQ Sequence 3300 BP; 845 A; 878 C; 877 G; 700 T; 0 U; 0 Other;

Query Match 97.5%; Score 3216.8; DB 4; Length 3300;

Best Local Similarity 98.4%; Pred. No. 0;

Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 1 CGGACACCATCGAATGGTGCAAAACCTTTTCGGGGTATGGCATGATAGCGCCCGGAAGAGA 60

Db 1 CGGACACCATCGAATGGTGCAAAACCTTTTCGGGGTATGGCATGATAGCGCCCGGAAGAGA 60

Qy 61 GTCAATTGAGGTGTAATGTGAACCAACGTAACGTTATACGATTCGAGAGTATGCGG 120

Db 61 GTCAATTGAGGTGTAATGTGAACCAACGTAACGTTATACGATTCGAGAGTATGCGG 120

Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCCAGCCACCGTTTCTCGGAAA 180

Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCCAGCCACCGTTTCTCGGAAA 180

Qy 181 CGCGGAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTATCATTTCCAAACCGCGTGGCAC 240

Db 181 CGCGGAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTATCATTTCCAAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGGTTGCCACCTCCAGTCTGGGCCCTTGC 300

Db 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGGTTGCCACCTCCAGTCTGGGCCCTTGC 300

Qy 301 ACGCCCGTCGCAAAATTGTGCGGCGGATTAAATCTCGCGCGATCAATCTGGGTGCGCAGCG 360

Db 301 ACGCCCGTCGCAAAATTGTGCGGCGGATTAAATCTCGCGCGATCAATCTGGGTGCGCAGCG 360

Qy 361 TGGTGGTGTGATGGTAGAACGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCAACAATC 420

Db 361 TGGTGGTGTGATGGTAGAACGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCAACAATC 420

Qy 421 TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGATCAACCAAGGATGCCA 480

Db 421 TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGATCAACCAAGGATGCCA 480

Qy 481 TTGCTGTGGAAGCTCGCTGCACTAATGTTCGGGGTTAATTTCTTGATGTCTCTGACCAAGA 540

Db 481 TTGCTGTGGAAGCTCGCTGCACTAATGTTCGGGGTTAATTTCTTGATGTCTCTGACCAAGA 540

Qy 541 CACCATCAACAGTATTAATTTTCTCCATGAAGACGGTACGGACTGGCGCTGGAGCATC 600

Db 541 CACCATCAACAGTATTAATTTTCTCCATGAAGACGGTACGGACTGGCGCTGGAGCATC 600

Qy 601 TGGTGGCATTCGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660

Db 601 TGGTGGCATTCGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660

Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Qy 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTTTCACCAACCAATGCAAAATGCTGA 780

Db 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTTTCACCAACCAATGCAAAATGCTGA 780

Qy 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGGTCAACAGATCAGATGCGCTGGCGGCAA 840

Db 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGGTCAACAGATCAGATGCGCTGGCGGCAA 840

Qy 841 TCGCGCCATTACCGAGTCCGGGCTGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900

Db 841 TCGCGCCATTACCGAGTCCGGGCTGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900

Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAACAGGATTTTC 960

Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAACAGGATTTTC 960

Qy 961 GCCTGCTGGGCAACAGCGCTGGACCGCTTGTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020

Db 961 GCCTGCTGGGCAACAGCGCTGGACCGCTTGTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020

Qy 1021 AGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080

Db 1021 AGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGAGTGGTGGTGGTGGTGGTGGT 1140

Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGAGTGGTGGTGGTGGTGGTGGT 1140

Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGAGTGGTGGTGGTGGTGGTGGT 1200

Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGAGTGGTGGTGGTGGTGGTGGT 1200

Qy 1201 GCACAATTCATGTTTTCAGACGCTTATCATCGATGCAAGCGGTGCAACCAATGCTTCTGGCG 1260

Db 1201 GCACAATTCATGTTTTCAGACGCTTATCATCGATGCAAGCGGTGCAACCAATGCTTCTGGCG 1260

QY	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCTGTAATCACTGATGAAGGCTG	1320
Db	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCTGTAATCACTGATGAAGGCTG	1320
QY	1321	TGTCGCTCAAGCGCAGCTCCCGTCTGATATGTTTTTGGCCGAGCATATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCAGCTCCCGTCTGATATGTTTTTGGCCGAGCATATACGGTT	1380
QY	1381	CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
Db	1381	CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
QY	1441	ATTGTAGGGGATACAAATTTACACAGAAACAGCCAGTCGTTAGGTGTTTCACGA	1500
Db	1441	ATTGTAGGGGATACAAATTTACACAGAAACAGCCAGTCGTTAGGTGTTTCACGA	1500
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGGAAGGTAAACTGGTAACTCTGG	1560
Db	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGGAAGGTAAACTGGTAACTCTGG	1560
QY	1561	ATTAAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT	1620
Db	1561	ATTAAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT	1620
QY	1621	ACCGGAATTAAGTCACCGTTGAGCATCCGGATAACTGGAAGAAATTCGCCAGGTT	1680
Db	1621	ACCGGAATTAAGTCACCGTTGAGCATCCGGATAACTGGAAGAAATTCGCCAGGTT	1680
QY	1681	CGCGCAACTGCGGATGGCCCTGACATTAATCTTGGGCACACGCGTTGGTGGTAC	1740
Db	1681	CGCGCAACTGCGGATGGCCCTGACATTAATCTTGGGCACACGCGTTGGTGGTAC	1740
QY	1741	GCTCAATCTGCGCTGTGGCTGAAATCAACCGGACAAAGGTTCCAGGACAAAGCTGTAT	1800
Db	1741	GCTCAATCTGCGCTGTGGCTGAAATCAACCGGACAAAGGTTCCAGGACAAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGATGCGTACGTTACAAAGGAAAGTAAAGCGGCTGATGCTTACCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGATGCGTACGTTACAAAGGAAAGTAAAGCGGCTGATGCTTACCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCGGAACCCGCAAAACCTCGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCGGAACCCGCAAAACCTCGGAA	1920
QY	1921	GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGCGGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGCGGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTTCACCTGCGCGCTGATGCTGCTGACGGGGTTATGCTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTTCACCTGCGCGCTGATGCTGCTGACGGGGTTATGCTTCAAG	2040
QY	2041	TATGAAACCGCAAGTACGATTAAGAACTGCGGCGTGGATACGCTGCGCGGAAAGCG	2100
Db	2041	TATGAAACCGCAAGTACGATTAAGAACTGCGGCGTGGATACGCTGCGCGGAAAGCG	2100
QY	2101	GCTGACCTTCTGTTGACCTGATTAAGAACTGAAAGCGAAAGTAAAGCGGCTGATGTTCAAC	2160
Db	2101	GCTGACCTTCTGTTGACCTGATTAAGAACTGAAAGCGAAAGTAAAGCGGCTGATGTTCAAC	2160
QY	2161	TCCATCGCAGAGTGCCTTTTAAAGAGCGAAACAGCGATGACCATCAACGCCCGTGG	2220
Db	2161	TCCATCGCAGAGTGCCTTTTAAAGAGCGAAACAGCGATGACCATCAACGCCCGTGG	2220
QY	2221	GCATGTCCAACTCGACACACGAAAGTGAATTTATGTTAAGCGTACTGCGGACCTTC	2280
Db	2221	GCATGTCCAACTCGACACACGAAAGTGAATTTATGTTAAGCGTACTGCGGACCTTC	2280
QY	2281	AAGGTCACACCTCAACCGTCTGTTGGCTGCTGAGCGCAGGTATTAACGGCGGCGAGT	2340
Db	2281	AAGGTCACACCTCAACCGTCTGTTGGCTGCTGAGCGCAGGTATTAACGGCGGCGAGT	2340
QY	2341	CCGAACAAGAGCTGGCAAAAGAGTTCTCCGAAAACTATCTGCTGACTGATGAAGGCTG	2400

Db	2341	CCGAACAAGAGCTGGCAAAAGAGTTCTCCGAAAACTATCTGCTGACTGATGAAGGCTG	2400
QY	2401	GAAGCGGTTAATAAGACAAACCGCTGGTCCCGTAGCTGAAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGTCCCGTAGCTGAAAGTCTTACGAGGAAGAG	2460
QY	2461	TTGGCGAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG	2520
QY	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCCGCTGCTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCCGCTGCTACTGCGGTGATCAACGCC	2580
QY	2581	GCACGCGTCTGACAGTCTGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG	2640
Db	2581	GCACGCGTCTGACAGTCTGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG	2640
QY	2641	AACAAACAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	2700
Db	2641	AACAAACAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	2700
QY	2701	GGATCTCTTCTCTGTCGCGCGGCTCTGAGCCCGGGGAGAGCCCTATGCTTGT	2760
Db	2701	GGATCTCTTCTCTGTCGCGCGGCTCTGAGCCCGGGGAGAGCCCTATGCTTGT	2760
QY	2761	CCGGATGTCGTAAGTCTTTCAGCGAGAGAGCTCTGTCGCGGAGAGAGAGAGAGAGAG	2820
Db	2761	CCGGATGTCGTAAGTCTTTCAGCGAGAGAGCTCTGTCGCGGAGAGAGAGAGAGAGAG	2820
QY	2821	ACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCCAGTCCAGCAAC	2880
QY	2881	CTGGTGGCGCATCAACGCACTACTTGGCGAGAGCCATACAAATGTCAGAAATGTGGC	2940
Db	2881	CTGGTGGCGCATCAACGCACTACTTGGCGAGAGCCATACAAATGTCAGAAATGTGGC	2940
QY	2941	AAGTCCTTTCAGCCAGTCCAGCACTGTCGCGCCACCAAGCTACTCACACGGGGAGAAG	3000
Db	2941	AAGTCCTTTCAGCCAGTCCAGCACTGTCGCGCCACCAAGCTACTCACACGGGGAGAAG	3000
QY	3001	CCCTATGTCCTGTCGCGGAAATGTTAAAGTCTTTCAGCACCAAGTGGCTCTTGGTTAGACAC	3060
Db	3001	CCCTATGTCCTGTCGCGGAAATGTTAAAGTCTTTCAGCACCAAGTGGCTCTTGGTTAGACAC	3060
QY	3061	CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGT	3120
Db	3061	CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGT	3120
QY	3121	CAGCGCGCCACCTGGAACGCGCATCAACGCACTCTACTTGGCGGAGAGCCATCAAAATGT	3180
Db	3121	CAGCGCGCCACCTGGAACGCGCATCAACGCACTCTACTTGGCGGAGAGCCATCAAAATGT	3180
QY	3181	CCAGAAATGTGGCAAGTCTTCTCAACTTCAGGCACTTCCGTCGCTCAGCAACGCTACTCAC	3240
Db	3181	CCAGAAATGTGGCAAGTCTTCTCAACTTCAGGCACTTCCGTCGCTCAGCAACGCTACTCAC	3240
QY	3241	ACCGGTAAAAAACTAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3300
Db	3241	ACCGGTAAAAAACTAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3300

RESULT 6

AH25681
ID AH25681 standard; DNA; 8101 BP.
XX
AC AH25681;
XX
DT 05-SEP-2001 (first entry)
XX
DE Nucleotide sequence of human Cdc25C phosphatase DNA.

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
XX ss.
XX
OS Homo sapiens.
XX
PN WO200144467-A2.
XX
PD 21-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-FR003496.
XX
PR 14-DEC-1999; 99FR-00015722.
PR 30-MAY-2000; 2000FR-00006883.
PR 21-SEP-2000; 2000FR-00012008.
XX
PA (SCRS) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
PI Goubin-Gramatica F, Ducommun B, Prevost G;
XX
DR WPI; 2001-398152/42.
XX
PT New fusion protein of maltose-binding protein and Cdc phosphatase, useful
PT for identifying phosphatase modulators for regulating the cell cycle.
XX
PS Claim 7; Page 9-13; 56pp; French.
XX
CC The present sequence encodes human Cdc25C phosphatase, and was expressed
CC in Escherichia coli. It is used to make fusion proteins with Escherichia
CC coli maltose binding protein (MBP). The specification describes fusion
CC proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
CC are involved in activation of cyclin-dependent kinases (CDK) that control
CC cell division. The fusion proteins are used to identify modulators of the
CC specified human Cdc phosphatases, potentially useful for regulation of
CC cell division. They are also for studying physiological or
CC physiopathological activities of Cdc phosphatase
XX
SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;
Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGTGAATGTGAACCAAGTAACTTATACGATGTCCGAGATATGCG 120
DB 61 GTCATTCAGGGTGTGAATGTGAACCAAGTAACTTATACGATGTCCGAGATATGCG 120
QY 121 GTGTCTTATCAGACCGTTTCCCGGTGTGAACCAAGTAACTTATACGATGTCCGAGAAA 180
DB 121 GTGTCTTATCAGACCGTTTCCCGGTGTGAACCAAGTAACTTATACGATGTCCGAGAAA 180
QY 181 CCGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
DB 181 CCGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCGACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCGACCTCCAGTCTGGCCCTGC 300
QY 301 AGCGCCGTTCGCAAAATGTTCGGCGGATTTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
DB 301 AGCGCCGTTCGCAAAATGTTCGGCGGATTTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTAGAACGAAGCGCGTTCGAAGCCCTGTAAAGCGCGGTGCACAATC 420
DB 361 TGGTGGTGTGATGTAGAACGAAGCGCGTTCGAAGCCCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTAGTGGGCTGATCAATTAATCTCCGCTGGATGACAGGATGCCA 480

DB 421 TTCTCGCGCAACCGCTAGTGGGCTGATCAATTAATCTCCGCTGGATGCCAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATCTCTGTGACCA 540
DB 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATCTCTGTGACCA 540
QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGAGGTACCGGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGAGGTACCGGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCAATTCGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
DB 601 TGGTCGCAATTCGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTCAACAAACCATGCAATGCTGA 780
DB 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGCA 840
DB 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGCA 840
QY 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATAG 900
DB 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATAG 900
QY 901 AGCATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACACCATCAACAGGATTTTC 960
DB 901 AGCATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGTGA 1020
DB 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGTGA 1020
QY 1021 AGGCAATCAGCTGTGTCCTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTGTCCTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCGAGTGGCAACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCGAGTGGCAACAGGTTT 1140
QY 1141 CCCGACTGGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTCTACTCATTTAG 1200
DB 1141 CCCGACTGGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTCTACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACCCATCGGAAGCTGTGATGCGTGTGAGGTGTAATCACTGCGATTAATTCG 1320
DB 1261 TCAGGACCCATCGGAAGCTGTGATGCGTGTGAGGTGTAATCACTGCGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGGAGATCATACGTT 1380
DB 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGGAGATCATACGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGA 1440
QY 1441 ATTGTGACGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
DB 1441 ATTGTGACGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
QY 1501 GCATTTCAACAGGAGGACCATAGATTATGAACCTGAAGAGGTAAATCTGGAATCTCG 1560
DB 1501 GCATTTCAACAGGAGGACCATAGATTATGAACCTGAAGAGGTAAATCTGGAATCTCG 1560

Db 1501 GCACCTTCCACCAACAGGACCATTAGATTATGAAAAATCGAAGAGTAAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAAGTTCGGAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAAGTTCGGAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGT 1680
Qy 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTAGCTTTACACGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTTACACGCAAGCTGATGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTCGCAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTCGCAACCCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACGCTACTTACCTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTACCTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAACAACAGCGATGACATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAACAACAGCGATGACATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACCAAGTGAATTTATGGTGAACGGTACTGSCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCAAGTGAATTTATGGTGAACGGTACTGSCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAAAACCGCTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGCTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGATGATGAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGATGATGAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2460
Qy 2461 TTGGCGAAAGATCCACGATTTGCGCCCAACCATGAAACCGCCGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGCCCAACCATGAAACCGCCGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGGTGGTATGCTGGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGGTGGTATGCTGGGTGATCAAGCC 2580
Qy 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCG 2640
Db 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCG 2640

Qy 2641 AACCAACAACAATAACAATAACAACAACCTCGGATCGAGGAAAGATTTCAGAAATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACAACCTCGGATCGAGGAAAGATTTCAGAAATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 7
ABL53238
ID ABL53238 standard; DNA; 8101 BP.
XX
AC ABL53238;
XX
DT 17-JUN-2002 (first entry)
XX
DE Nucleotide sequence of pMAL-Hs Cdc25C.
XX
KW Cdc25C phosphatase; amine; Cdc25 phosphatase; spontaneous alopecia;
KW alopecia; proliferative disease; parasitic disease; viral infection;
KW neurodegenerative disease; myopathy; angiogenesis; psoriasis; restenosis;
KW vitamin K; SA.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key
FT CDS Location/Qualifiers
FT FT /*tag= b
FT FT /note= "encodes maltose binding protein (MBP)"
FT FT 2713..4134
FT FT /*tag= a
FT FT /note= "ORF of human Cdc25C"
XX
FN FR2812198-A1.
XX
PD 01-FEB-2002.
XX
PF 28-JUL-2000; 2000FR-00009900.
XX
PR 28-JUL-2000; 2000FR-00009900.
XX
PA (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
PI Prevost G, Brezak PMC, Galcera CMO, Thureau C, Goubin GF;
PI Ducommun B;
XX
WP1; 2002-166066/22.
XX
PT cdc25 Phosphatase inhibiting medicaments, useful e.g. for treating tumor
PT diseases, viral infections, neurodegenerative disease or alopecia,
PT containing new or known aryl-substituted amines.
XX
PS Disclosure; Page 18-22; 39pp; French.
XX
CC The present sequence represents pMAL-HsCdc25C, a plasmid encoding human
CC Cdc25C phosphatase. The specification describes the use of aryl-
CC substituted secondary or tertiary amines for the production of
CC medicaments for inhibiting Cdc25 phosphatases, especially Cdc25C
CC phosphatases. The amines of the invention are used for treating
CC spontaneous alopecia or alopecia induced by exogenous products or
CC radiation. They are also used for treating tumoral or non-tumoral
CC proliferative diseases, parasitic diseases, viral infections,
CC neurodegenerative diseases or myopathy. Examples of non-tumoral
CC proliferative diseases are e.g. angiogenesis, psoriasis or restenosis.
CC The amines are also useful for inhibiting the proliferation of
CC microorganisms (especially yeasts) and in the treatment of all disorders
CC conventionally treated with vitamin K or its derivatives
XX
SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match 82.0%; Score 2705.8; DB 6; Length 8101;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	1	CGGACACCATCGAATGGGTGCAAAACCTTTCGGCGGTATGGCATGATAGGCCCGGGAAGAGA	60						
DB	1	CGGACACCATCGAATGGGTGCAAAACCTTTCGGCGGTATGGCATGATAGGCCCGGGAAGAGA	60						
QY	61	GTCAATTCAGGTGTGTAATGTGAACACAGTAAACGTTATACGATGTGCGAGAGTATGCGG	120						
DB	61	GTCAATTCAGGTGTGTAATGTGAACACAGTAAACGTTATACGATGTGCGAGAGTATGCGG	120						
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGTGAACACGCGCCAGCCACGTTTCTGCGAANA	180						
DB	121	GTGTCCTTATCAGACCGTTTCCCGGTGTGAACACGCGCCAGCCACGTTTCTGCGAANA	180						
QY	181	CGCGGAAAGAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCAACCGCTGGGCAC	240						
DB	181	CGCGGAAAGAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCAACCGCTGGGCAC	240						
QY	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
DB	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
QY	301	ACGCGCGCTCGCAAAATTTGTCGGGGGATTAATCTCGGCCGATCAATCTGGGTGCCAGG	360						
DB	301	ACGCGCGCTCGCAAAATTTGTCGGGGGATTAATCTCGGCCGATCAATCTGGGTGCCAGG	360						
QY	361	TGTCGTGTCGATGGTAGAACGAGCGGCTGCAAGCCCTGTAAGCGGGGTGCAACAATC	420						
DB	361	TGTCGTGTCGATGGTAGAACGAGCGGCTGCAAGCCCTGTAAGCGGGGTGCAACAATC	420						
QY	421	TTCTCGCGCAACGCTGAGTGGGTGATCAATTAATCTCCGCTGGATGCCAGGATGCCA	480						
DB	421	TTCTCGCGCAACGCTGAGTGGGTGATCAATTAATCTCCGCTGGATGCCAGGATGCCA	480						
QY	481	TTGCTGTGGAGCTGCCTGCATTAATGTTCCGGCGTTATTTCTGATGTCCTGACCCAGA	540						
DB	481	TTGCTGTGGAGCTGCCTGCATTAATGTTCCGGCGTTATTTCTGATGTCCTGACCCAGA	540						
QY	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
DB	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
QY	601	TGGTCGCATTTGGGTCAACGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCCTGTCGG	660						
DB	601	TGGTCGCATTTGGGTCAACGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCCTGTCGG	660						
QY	661	CGCGTCTGGCTGCGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
DB	661	CGCGTCTGGCTGCGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
QY	721	CGGAAACGGAGCGGACGTGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780						
DB	721	CGGAAACGGAGCGGACGTGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780						
QY	781	ATGAGGGCATCGTTCCTCCATCTGCGATGCTGTTTGCCTGCAACGATCAGATGGCGTGGCGCA	840						
DB	781	ATGAGGGCATCGTTCCTCCATCTGCGATGCTGTTTGCCTGCAACGATCAGATGGCGTGGCGCA	840						
QY	841	TGCGGCGCATACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAGC	900						
DB	841	TGCGGCGCATACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAGC	900						
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAACAGGATTTTC	960						
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAACAGGATTTTC	960						
QY	961	GCCTGTGGGGCAAAACCGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020						
DB	961	GCCTGTGGGGCAAAACCGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020						
QY	1021	AGGGCAATCAGCTGTTGCCCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA	1080						

DB	1021	AGGGCAATCAGCTGTTGCCCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA	1080						
QY	1081	CGCAACCGCCTCTCTCCCGCGCGTTCGCGGATTCATTAAATGCAGTGGCAGCACAGGTTT	1140						
DB	1081	CGCAACCGCCTCTCTCCCGCGCGTTCGCGGATTCATTAAATGCAGTGGCAGCACAGGTTT	1140						
QY	1141	CCGACTGGAAGCGGCGAGTGAGGCAACGGAATTAATGTGAGTTAGTCTACTCATTTAG	1200						
DB	1141	CCGACTGGAAGCGGCGAGTGAGGCAACGGAATTAATGTGAGTTAGTCTACTCATTTAG	1200						
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATGCAGTGCAGGTGCACCAATCTTCTGGCG	1260						
DB	1201	GCACAATTTCTCATGTTTGACAGCTTATCATGCAGTGCAGGTGCACCAATCTTCTGGCG	1260						
QY	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCTGTAAATCACTGCATAAATTCG	1320						
DB	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCTGTAAATCACTGCATAAATTCG	1320						
QY	1321	TGTGCTCAAGCGGCACTCCCGTTCGATTAATGTTTTTGGCGCGGACATCAATACGGTT	1380						
DB	1321	TGTGCTCAAGCGGCACTCCCGTTCGATTAATGTTTTTGGCGCGGACATCAATACGGTT	1380						
QY	1381	CTGSCAATATTTCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATATGTGTGA	1440						
DB	1381	CTGSCAATATTTCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATATGTGTGA	1440						
QY	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAACGCCAGTCCGTTTAGTGTTCACGA	1500						
DB	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAACGCCAGTCCGTTTAGTGTTCACGA	1500						
QY	1501	GCATTTCACCAAGGACCATAGATTATGAAATCGAAGAAAGGTAAATCTGGTAATCTGG	1560						
DB	1501	GCATTTCACCAAGGACCATAGATTATGAAATCGAAGAAAGGTAAATCTGGTAATCTGG	1560						
QY	1561	ATTAACGGCGATAAAGGCTATAACGGCTCCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620						
DB	1561	ATTAACGGCGATAAAGGCTATAACGGCTCCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620						
QY	1621	ACCGGAATTTAAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGCCAGGTT	1680						
DB	1621	ACCGGAATTTAAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGCCAGGTT	1680						
QY	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCAACGACCGCTTTGGTGGCTAC	1740						
DB	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCAACGACCGCTTTGGTGGCTAC	1740						
QY	1741	GCTCAATCTGGCCTGTTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800						
DB	1741	GCTCAATCTGGCCTGTTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800						
QY	1801	CCGTTTACCTGGATGCGGTACGTTTCAACCGCAAGCTGATTTGCTTACCCGATCGCTGTT	1860						
DB	1801	CCGTTTACCTGGATGCGGTACGTTTCAACCGCAAGCTGATTTGCTTACCCGATCGCTGTT	1860						
QY	1861	GAAGCGTTATCGCTGATTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
DB	1861	GAAGCGTTATCGCTGATTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAGGTGAAGCGCGCTGATGTTCAAC	1980						
DB	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAGGTGAAGCGCGCTGATGTTCAAC	1980						
QY	1981	CTGCAAGAACCGTACTCTCACCTGGCGCTGATTTGCTGTGACGGGGTTATCGCTTCAAG	2040						
DB	1981	CTGCAAGAACCGTACTCTCACCTGGCGCTGATTTGCTGTGACGGGGTTATCGCTTCAAG	2040						
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACCGTGGCGTGGATTAACGCTGGCGCGAAAGCG	2100						
DB	2041	TATGAAAAACGCAAGTACGACATTAAGACCGTGGCGTGGATTAACGCTGGCGCGAAAGCG	2100						
QY	2101	GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAAACACATGAATGCGAGACCGGATAC	2160						

```
Db 2101 GGTCTGACCTTCTCGTGGTACCTGATTAAAAACAACACATGATGATGACACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGCAACAGCGATGACCATCAACGCGCGGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGCAACAGCGATGACCATCAACGCGCGGTGG 2220
QY 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATTATGTGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATTATGTGTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAACACCGTTCGTTGGCGTGTGAGGCGAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAACACCGTTCGTTGGCGTGTGAGGCGAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAACACCGCTGGGTGCGGTGAGCTGAAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAACACCGCTGGGTGCGGTGAGCTGAAAGTCTTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGATTATGCGCCACCATGGAACACGCGCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGATTATGCGCCACCATGGAACACGCGCAGAAAGGTGAATCATG 2520
QY 2521 CCGAATCCCGCAGATGTCGCGCTTTCTGTTGATGCGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAATCCCGCAGATGTCGCGCTTTCTGTTGATGCGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATCGAGTCTG 2640
QY 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGGATTTTCAAGATTC 2700
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709
```

RESULT 8
ADO23608

```
ID ADO23608 standard; DNA; 7259 BP.
AC
XX
AC ADO23608;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30del18(82-294aa) fusion protein.
XX
KW P30 antigen; Toxop30del13C; Toxop30del12C; Toxop30 MIX1;
KW MBP-Toxop30del13C(52-300aa); MBP-Toxop30del14C(52-294aa); MBP-Toxop30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
FN US2004067239-A1.
XX
XX 08-APR-2004.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX (MAIN/) MAINE G T.
XX (PATE/) PATEL C B.
XX (GINS/) GINSBURG S R.
XX (BLIE/) BLIESE T R.
```

```
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23609.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxop30del13C,
PT Toxop30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
XX Example 2; Fig 22; 114pp; English.
XX
XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxop30del13C, Toxop30del12C and Toxop30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxop30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxop30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxop30del13C(52-300aa), MBP-Toxop30del14C(52-294aa) and MBP-Toxop30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
XX
```

Sequence 7259 BP; 1870 A; 1822 C; 1926 G; 1641 T; 0 U; 0 Other;

```
Query Match 81.5%; Score 2690.4; DB 12; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCACACCATCGAATGCTGCAAAACCTTTTCGCGGTATGGCATGATAGCCCCGGAAGAGA 60
Db 1 CCACACCATCGAATGCTGCAAAACCTTTTCGCGGTATGGCATGATAGCCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACACCAAGTACGTTATACGATGTCGCAGATATGCCG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAACACCAAGTACGTTATACGATGTCGCAGATATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCCGAAA 180
QY 181 CGCGGGAAGAGTGGAGCGGCGATGCGCGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Db 181 CGCGGGAAGAGTGGAGCGGCGATGCGCGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACACAGTTCGTTCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACACAGTTCGTTCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 AGCGCGCGTCGCAAAATGTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 AGCGCGCGTCGCAAAATGTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGTCGATGTTAGAACGAGCGGCTCGAAGCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGGTGTGTGTCGATGTTAGAACGAGCGGCTCGAAGCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCGTCAGTGGGCTGATCATTAATATATCCGCTGGATGACACGAGTGCCA 480
Db 421 TTCTCGCGCAACCGCGTCAGTGGGCTGATCATTAATATATCCGCTGGATGACACGAGTGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGCGCTTATTTCTTGATGTCCTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGCGCTTATTTCTTGATGTCCTGACACAGA 540
QY 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGGTGGAGCATC 600
```

QY 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCCTCGG 660
DB 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCCTCGG 660
QY 661 CGCGTCTCGCTGCGCTGGCGTAAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTCGCTGCGCTGGCGTAAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTCAACAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTCAACAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTCACACGATCAGATGCGCTGGGCGCAA 840
DB 781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTCACACGATCAGATGCGCTGGGCGCAA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
DB 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCGAAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCGAAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAACACACCTGGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAACACACCTGGGCGCCCAATA 1080
QY 1081 CGCAAACCGCCTCTCCCGCGGTTGGCCGATTCAATTAATGCGTGGCAGCAGGTTT 1140
DB 1081 CGCAAACCGCCTCTCCCGCGGTTGGCCGATTCAATTAATGCGTGGCAGCAGGTTT 1140
QY 1141 CCGGACTCGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTACTACTCATTTAG 1200
DB 1141 CCGGACTCGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTACTACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGTGGTGCACCAATGTTCTGGCG 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGTGGTGCACCAATGTTCTGGCG 1260
QY 1261 TCAGCAGCCATTCGAAGCTGGTATGCTGTGAGGTTCGTAATCACTGCATAAATTCG 1320
DB 1261 TCAGCAGCCATTCGAAGCTGGTATGCTGTGAGGTTCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
DB 1321 TGTGCTCAAGCGCACTCCCGTTCGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTGATATGTTGTA 1440
DB 1381 CTGGCAAAATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTGATATGTTGTA 1440
QY 1441 ATTGTGAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGTGTTTTCAGA 1500
DB 1441 ATTGTGAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGTGTTTTCAGA 1500
QY 1501 GCACCTTACCAACAAAGGACATAGATTATGAAAACTGAAGAAAGGTAATCTGGTAATCTGG 1560
DB 1501 GCACCTTACCAACAAAGGACATAGATTATGAAAACTGAAGAAAGGTAATCTGGTAATCTGG 1560
QY 1561 ATTAACGCGATAAAGGCTATAACGCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
DB 1561 ATTAACGCGATAAAGGCTATAACGCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740

DB 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGGACAAGAGGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGGACAAGAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTCGTTACCCGATCCGCTGT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTCGTTACCCGATCCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTATAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGCAAGTACGACATTAAAGACGCTGGGCGTGGATAACGCTGGGCGGAAAGCG 2100
DB 2041 TATGAAAACGCAAGTACGACATTAAAGACGCTGGGCGTGGATAACGCTGGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAAACACATGAAATGCGAGACCCGATTAC 2160
DB 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAAACACATGAAATGCGAGACCCGATTAC 2160
QY 2161 TCCATCGAGAAGCTGCTTTAAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
DB 2161 TCCATCGAGAAGCTGCTTTAAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACACGAGCAAAAGTGAATTAATGCTGTAACGCTGCGGACCTTC 2280
DB 2221 GCATGGTCCAAATCGACACACGAGCAAAAGTGAATTAATGCTGTAACGCTGCGGACCTTC 2280
QY 2281 AAGGTCACACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
DB 2281 AAGGTCACACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
DB 2341 CCGAACAAGAGCTGGCAAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGTTTAATAAGACAAACCGCTGGTGGCTGAGCGCTGAACTTTACGAGGAAG 2460
DB 2401 GAAGCGTTTAATAAGACAAACCGCTGGTGGCTGAGCGCTGAACTTTACGAGGAAG 2460
QY 2461 TTGGCAAGAGTCCACGTTATTTGGCGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGGCAAGAGTCCACGTTATTTGGCGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACAATCCCGCAGATCCCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
DB 2521 CCGAACAATCCCGCAGATCCCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAGCCCTGGAAGACGCGCAGACTTAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGACAGCTGTCGATGAGCCCTGGAAGACGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAACAATAACAACAACTCGGGAATCGAGGGAAGGATTTCAGAAATTC 2700
DB 2641 AACCAACAACAATAACAATAACAACAACTCGGGAATCGAGGGAAGGATTTCAGAAATTC 2700

RESULT 9
ADO23613
ID ADO23613 standard; DNA; 7322 BP.
XX
AC ADO23613;

01-JUL-2004 (first entry)
 DNA encoding MBP-Toxop30del110(52-284aa) fusion protein.
 P30 antigen; Toxop30del13C; Toxop30del12C; Toxop30 MIX1;
 MBP-Toxop30del13C(52-300aa); MBP-Toxop30del14C(52-294aa); MBP-Toxop30MIX1;
 Toxoplasma gondii; ds; gene.
 Toxoplasma gondii.
 Synthetic.
 US2004067239-A1.
 08-APR-2004.
 02-OCT-2002; 2002US-00263153.
 02-OCT-2002; 2002US-00263153.
 (MAIN/) MAINE G T.
 (PATE/) PATEL C B.
 (GINS/) GINSBURG S R.
 (BLIE/) BLIESE T R.
 Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 WPI; 2004-304563/28.
 P-PSDB; ADO23614.
 Novel purified polypeptide having sequence identity to amino acid
 sequence of P30 antigen of Toxoplasma gondii e.g. Toxop30del13C,
 Toxop30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
 antibodies to Toxoplasma gondii.
 Example 2; Fig 25; 114pp; English.
 The invention relates to a purified P30 antigen (I) chosen from 3 fully
 defined Toxop30del13C, Toxop30del12C and Toxop30 MIX1 sequences, having 1-6
 amino acids added to the C-terminus of the amino acid sequence of
 Toxop30del12C P30 antigen sequence, or comprising an amino acid sequence in
 which at least one of the five C-terminal cysteine amino acids of the
 amino acid sequence of Toxop30del13C P30 antigen sequence is substituted
 with alanine, or comprising the amino acid sequence chosen from MBP-
 Toxop30del13C(52-300aa), MBP-Toxop30del14C(52-294aa) and MBP-Toxop30MIX1
 fusion proteins. (I) is useful for detecting the presence of IgM
 antibodies to Toxoplasma gondii in a test sample, which involves
 contacting the test sample suspected of containing the IgM antibodies
 with a composition comprising (I) and detecting the presence of (I)/IgM
 antibody complexes. The present sequence represents DNA encoding a MBP-
 Toxop30 fusion protein of the invention.
 Sequence 7322 BP; 1887 A; 1844 C; 1934 G; 1657 T; 0 U; 0 Other;
 Query Match 81.5%; Score 2690.4; DB 12; Length 7322;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 Qy 1 CCGACACCATCGAATGGTGAACCACTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
 Db |
 Qy 1 CCGACACCATCGAATGGTGAACCACTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
 Db |
 Qy 61 GTCAATTCAGGGTGGTGAATGTGAACCACTTAAGTTATACGATGTCGAGATATGCCG 120
 Db |
 Qy 61 GTCAATTCAGGGTGGTGAATGTGAACCACTTAAGTTATACGATGTCGAGATATGCCG 120
 Db |
 Qy 121 GTGTCTTTATCAGACCGTTTCCCGCGTGTGAACCACTTAAGTTATACGATGTCGAGATATGCCG 180
 Db |
 Qy 121 GTGTCTTTATCAGACCGTTTCCCGCGTGTGAACCACTTAAGTTATACGATGTCGAGATATGCCG 180
 Db |
 Qy 181 CGCGGGAAGTGAACCGCGGATGCGGAGCTGAATTCATTCCTCCAAACCGCGTGGCAC 240
 Db |
 Qy 181 CGCGGGAAGTGAACCGCGGATGCGGAGCTGAATTCATTCCTCCAAACCGCGTGGCAC 240
 Db |

241 AACAACTGGCGGGCAAAACAGTCGTTGTGATTTGGCGTTGCACCTCCAGTCTGCGCCCTGC 300
 Db |
 241 AACAACTGGCGGGCAAAACAGTCGTTGTGATTTGGCGTTGCACCTCCAGTCTGCGCCCTGC 300
 Qy |
 301 ACGCGCGTTCGCAAAATTTGTCGCGCGGATTAATCTTCGCGCGGATCAACTGGTGTCCAGCG 360
 Db |
 301 ACGCGCGTTCGCAAAATTTGTCGCGCGGATTAATCTTCGCGCGGATCAACTGGTGTCCAGCG 360
 Qy |
 361 TGGTGGTGTGATGTTAGAACGAAAGCGCGCTCGAAGCGCTGTAAAGCGCGGTGCACAATC 420
 Db |
 361 TGGTGGTGTGATGTTAGAACGAAAGCGCGCTCGAAGCGCTGTAAAGCGCGGTGCACAATC 420
 Qy |
 421 TTCTCGCGCAACGCGTCAGTGGCTGATCAATTAACATATCCGCTGGATGACAGGATGCCA 480
 Db |
 421 TTCTCGCGCAACGCGTCAGTGGCTGATCAATTAACATATCCGCTGGATGACAGGATGCCA 480
 Qy |
 481 TTGCTGTGGAAGCTGCGTCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
 Db |
 481 TTGCTGTGGAAGCTGCGTCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
 Qy |
 541 CACCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGCGTGGAGCATC 600
 Db |
 541 CACCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGCGTGGAGCATC 600
 Qy |
 601 TGGTCCGATTTGGGTCAACGAGCAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
 Db |
 601 TGGTCCGATTTGGGTCAACGAGCAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
 Qy |
 661 CGCGTCTGCGTCTGGCTGGCGTAAATATCTCACTCGCAATCAAAATCAAGTCTGACCATAG 720
 Db |
 661 CGCGTCTGCGTCTGGCTGGCGTAAATATCTCACTCGCAATCAAAATCAAGTCTGACCATAG 720
 Qy |
 721 CGGAAACGGGAAGCGGATGAGTGCATGTCGGTCTTCAACAAACCATGCAAAATGCTGA 780
 Db |
 721 CGGAAACGGGAAGCGGATGAGTGCATGTCGGTCTTCAACAAACCATGCAAAATGCTGA 780
 Qy |
 781 ATGAGGCGATCGTTTCCCACTCGCATCTGTTGCCAACCATGAGATGGCGTGGCGGCGAA 840
 Db |
 781 ATGAGGCGATCGTTTCCCACTCGCATCTGTTGCCAACCATGAGATGGCGTGGCGGCGAA 840
 Qy |
 841 TGGCGCGCATTTACCGAGTCCGGCTCGCGTGTGGTCCGATATCTCGTAGTGGGATACG 900
 Db |
 841 TGGCGCGCATTTACCGAGTCCGGCTCGCGTGTGGTCCGATATCTCGTAGTGGGATACG 900
 Qy |
 901 ACGATACCGAAGACACGCTCATGTTATATCCGCGGTTAACCCACCATCAACAGGATTTTC 960
 Db |
 901 ACGATACCGAAGACACGCTCATGTTATATCCGCGGTTAACCCACCATCAACAGGATTTTC 960
 Qy |
 961 GCCTGTGGGGCAAAACGAGTGGACCGCTGTGCAACTCTCTCAGGGCCAGGGCGGTGA 1020
 Db |
 961 GCCTGTGGGGCAAAACGAGTGGACCGCTGTGCAACTCTCTCAGGGCCAGGGCGGTGA 1020
 Qy |
 1021 AGGGCAATCAGCTGTGTCCTGCTCACTGGTGAAGAAAAACCACTTGGCGGCCCAATA 1080
 Db |
 1021 AGGGCAATCAGCTGTGTCCTGCTCACTGGTGAAGAAAAACCACTTGGCGGCCCAATA 1080
 Qy |
 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTTCAATTAATGAGTGGCGGACGACAGGTTT 1140
 Db |
 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTTCAATTAATGAGTGGCGGACGACAGGTTT 1140
 Qy |
 1141 CCCGATGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG 1200
 Db |
 1141 CCCGATGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG 1200
 Qy |
 1201 GCACAAATTCATGTTTGAACGCTTATCATGCACTGCGGTGACCAATGCTTCTGGCG 1260
 Db |
 1201 GCACAAATTCATGTTTGAACGCTTATCATGCACTGCGGTGACCAATGCTTCTGGCG 1260
 Qy |
 1261 TCAGGACGACCATCGGAAGCTGTGGTATGCGTGTGAGGTCGTAATCACTGCAATATTCG 1320
 Db |
 1261 TCAGGACGACCATCGGAAGCTGTGGTATGCGTGTGAGGTCGTAATCACTGCAATATTCG 1320

QY 1321 TGTCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATATTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTACGA 1500
QY 1501 GCATTTCAACACAGGACCATAGATTATGAAAACTGAAGAGTTAAACTGTAATCTGG 1560
Db 1501 GCATTTCAACACAGGACCATAGATTATGAAAACTGAAGAGTTAAACTGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAATCGGTGAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAATCGGTGAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCCTCAAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCCTCAAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACAGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACAGCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGTGCTGAAATCAACCCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGCTGAAATCAACCCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGAAGCTGATTCCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGAAGCTGATTCCTTACCCGATCGCTGT 1860
QY 1861 GAAGGTTTATCGCTGATTTATACAAAGATCTGCTGCGAACCAGCGCAAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATACAAAGATCTGCTGCGAACCAGCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCTCGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTCGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGTGACGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGGTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGGTGGATAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATCAATGACGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATCAATGACGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACAGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCACTCAAAACCGTTGTTGGCTGCTGAGCGGAGGTTTAAACCGCCCGT 2340
Db 2281 AAGGTCACCACTCAAAACCGTTGTTGGCTGCTGAGCGGAGGTTTAAACCGCCCGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460

Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGCTATTGCGCCCAACCATGGAACACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATTGCGCCCACTATGGAACACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700

RESULT 10
ADO23603

ID ADO23603 standard; DNA; 7352 BP.

XX ADO23603;

XX 01-JUL-2004 (first entry)

XX DNA encoding MBP-ToxoP30del14C(52-294aa) fusion protein.

XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;

KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;

KW Toxoplasma gondii; ds; gene.

XX Toxoplasma gondii.

OS Synthetic.

XX US2004067239-A1.

XX 08-APR-2004.

XX 02-OCT-2002; 2002US-00263153.

XX 02-OCT-2002; 2002US-00263153.

XX (MAIN/) MAINE G T.

PA (PATE/) PATEL C B.

PA (GINS/) GINSBURG S R.

PA (BLIE/) BLIESE T R.

XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;

XX WPI; 2004-304563/28.

DR P-PSDB; ADO23604.

XX Novel purified polypeptide having sequence identity to amino acid sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C, Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG antibodies to Toxoplasma gondii.

XX Example 2; Fig 19; 114pp; English.

XX The invention relates to a purified P30 antigen (I) chosen from 3 fully defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6 amino acids added to the C-terminus of the amino acid sequence of Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in which at least one of the five C-terminal cysteine amino acids of the amino acid sequence of Toxo30del13C P30 antigen sequence is substituted with alanine, or comprising the amino acid sequence chosen from MBP-Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1 fusion proteins. (I) is useful for detecting the presence of IgM antibodies to Toxoplasma gondii in a test sample, which involves contacting the test sample suspected of containing the IgM antibodies

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
XX Toxop30 fusion protein of the invention.
SQ Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGGACACCATCAATGGTCAAAACCTTTCGGGTATGGCATGATAGGCCCGGAAGAGA	60
DB	1	CGGACACCATCAATGGTCAAAACCTTTCGGGTATGGCATGATAGGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG	120
DB	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG	120
QY	121	GGTCTCTTATCAGACCGTTCCTCCGGTGGTGAACACAGCCAGCCACGCTTTCGGAAAA	180
DB	121	GGTCTCTTATCAGACCGTTCCTCCGGTGGTGAACACAGCCAGCCACGCTTTCGGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTAATTCACATCCCAACCGCTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGGAGTGGCGGAGCTGAATTAATTCACATCCCAACCGCTGGCAC	240
QY	241	AACAACTGGCGGCAAAAGTGTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGGCGGCAAAAGTGTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTTCGCAATTCGCGGCGATTAATCTCGCGCGATCACTGGGTGCGACGC	360
DB	301	ACGCGCGTTCGCAATTCGCGGCGATTAATCTCGCGCGATCACTGGGTGCGACGC	360
QY	361	TGGTGGTTCGATGTAGAACGAAGCGCGTGAAGCGCTGAAGCGCGGTGCACAAATC	420
DB	361	TGGTGGTTCGATGTAGAACGAAGCGCGTGAAGCGCTGAAGCGCGGTGCACAAATC	420
QY	421	TTCTCGCGCAAGCGTCACTGAGTGGGTGATCAATTAATCTCCGCTGGATGACCAAGTCC	480
DB	421	TTCTCGCGCAAGCGTCACTGAGTGGGTGATCAATTAATCTCCGCTGGATGACCAAGTCC	480
QY	481	TTGCTGTGGAAGTCCCTGCAATTAATCTCCGCGGTATTTCTTGATGCTCTGACGAGA	540
DB	481	TTGCTGTGGAAGTCCCTGCAATTAATCTCCGCGGTATTTCTTGATGCTCTGACGAGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGCTGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGCTGGGCGTGGAGCATC	600
QY	601	TGGTCCGATTCGGTCAACGCAATTCGGCTGTTAGCGGCGCAATTAAGTTCCTCTCGG	660
DB	601	TGGTCCGATTCGGTCAACGCAATTCGGCTGTTAGCGGCGCAATTAAGTTCCTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGCGATGGAGTCCCATGTCGGTTCATCAACCAACCATGCAATGCTGA	780
DB	721	CGGAACGGGAAGGCGATGGAGTCCCATGTCGGTTCATCAACCAACCATGCAATGCTGA	780
QY	781	ATGAGGCGATCTTCCCACTCGCATGCTGGTTCGAACGATCAGATGGCGCTGGCGCAA	840
DB	781	ATGAGGCGATCTTCCCACTCGCATGCTGGTTCGAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGCGCATTCACGAGTCCGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGCGCATTCACGAGTCCGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAGACAGTCTATGTTATATCCGCGGTTAACCAATCAACAGGATTTTC	960
DB	901	ACGATACCGAGACAGTCTATGTTATATCCGCGGTTAACCAATCAACAGGATTTTC	960

QY	961	GCCTGCTGGGCAACACGAGCTGACCGCTTGCTGCAACTCTCTCAGGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGCAACACGAGCTGACCGCTTGCTGCAACTCTCTCAGGCCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGCCCCGTCTCACTGGTGAAGAAAAACACCTCGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTGCCCCGTCTCACTGGTGAAGAAAAACACCTCGCGCCCAATA	1080
QY	1081	CGAAAAACCGCTCTCCCGCGGTGGCGGATCAATTAATGACAGTGGCAGCAGGTTT	1140
DB	1081	CGAAAAACCGCTCTCCCGCGGTGGCGGATCAATTAATGACAGTGGCAGCAGGTTT	1140
QY	1141	CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGCTCATTAG	1200
DB	1141	CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGCTCATTAG	1200
QY	1201	GCACAAATCTCATGTTTTGACAGCTTATCATCGACTGACCGTGCACCAATGCTTCGCG	1260
DB	1201	GCACAAATCTCATGTTTTGACAGCTTATCATCGACTGACCGTGCACCAATGCTTCGCG	1260
QY	1261	TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATAATTCG	1320
DB	1261	TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATAATTCG	1320
QY	1321	TGTCGCTCAAGGCGCACTCCCGTTCGGAATAATGTTTTTCGCGCGACATCAACCGTT	1380
DB	1321	TGTCGCTCAAGGCGCACTCCCGTTCGGAATAATGTTTTTCGCGCGACATCAACCGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGTGGA	1440
QY	1441	ATTGTGAGCGGATCAAAATTTTCACAGGAAACAGCCAGTCCGTTAGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATCAAAATTTTCACAGGAAACAGCCAGTCCGTTAGTGTTCACGA	1500
QY	1501	GCATTTCCACCAACAGGACCATAGATTTATGAAATCGAAGAGTAACTGGTAAATCTCG	1560
DB	1501	GCATTTCCACCAACAGGACCATAGCATATGAAATCGAAGAGTAACTGGTAAATCTCG	1560
QY	1561	ATTAAACGGCGATAAAGCTATAAACGCTCTCGCTGAAAGTAAAGTAAAGTAAAGAT	1620
DB	1561	ATTAAACGGCGATAAAGGCTATAAACGCTCTCGCTGAAAGTAAAGTAAAGTAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACGATGGAAGAAATTCACAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACGATGGAAGAAATTCACAGGTT	1680
QY	1681	GGCGCAACTGGCGATGGCCCTGACATTTATCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
DB	1681	GGCGCAACTGGCGATGGCCCTGACATTTATCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCGTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGCAAGCTGTAT	1800
DB	1741	GCTCAATCTGGCCGTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGCAAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATCGCGTACGTTACAAACGCAAGCTGATTCGCTTACCGATCGCTGT	1860
DB	1801	CGGTTTACCTGGGATCGCGTACGTTACAAACGCAAGCTGATTCGCTTACCGATCGCTGT	1860
QY	1861	GAAGCGTTATCGCTGATTTAACAAGATCTGCTCCGAAACCCGCAAAAACTCTGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTAACAAGATCTGCTCCGAAACCCGCAAAAACTCTGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAGTAAAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATCGGTTCAAG	2040

```
QY 2041 TATGAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATTAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATTAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCCACATCGACACCGACGAACTGAATTAATGCTGAACGCTGCGGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACCGACGAACTGAATTAATGCTGAACGCTGCGGACCTTC 2280
QY 2281 AAGGCTCAACCATCAAAACCGTTCTGCGCGTCTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGCTCAACCATCAAAACCGTTCTGCGCGTCTGAGCGAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTCGCCCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTCGCCCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGAGCGCCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGAGCGCCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAA 2700
Db 2641 AACAAACAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAA 2700
```

RESULT 11

ADO23598

ID ADO23598 standard; DNA; 7370 BP.

XX AC ADO23598;

XX DT 01-JUL-2004 (first entry)

XX DE DNA encoding MBP-ToxoP30del13C(52-300aa) fusion protein.

XX KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
XX KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
XX KW Toxoplasma gondii; ds; gene.

XX OS Toxoplasma gondii.

XX OS Synthetic.

XX PN US2004067239-A1.

XX PD 08-APR-2004.

XX PF 02-OCT-2002; 2002US-00263153.

XX PR 02-OCT-2002; 2002US-00263153.

XX PA (MAIN/) MAINE G T.

XX PA (PATE/) PATEL C B.

XX PA (GINS/) GINSBURG S R.

(BLIE/) BLIESE T R.

XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;

XX DR WPI; 2004-304563/28.

XX DR P-PSDB; ADO23599.

XX PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.

XX PS Example 2; Fig 15; 114pp; English.

XX CC The invention relates to a purified p30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C p30 antigen sequence, or comprising an amino acid of the
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C p30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.

XX SQ Sequence 7370 BP; 1901 A; 1855 C; 1949 G; 1665 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGCGCATGATAGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGCGCATGATAGCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGGTGTGAATGTGAAACCAAGTAACTTATACGATCTCGCAGAGTATCGCG 120
Db 61 GTCAAATTCAGGGTGTGAATGTGAAACCAAGTAACTTATACGATCTCGCAGAGTATCGCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCTCCCGCGTGTGTAACACGAGCCAGCCAGCTTTCTCGAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCTCCCGCGTGTGTAACACGAGCCAGCCAGCTTTCTCGAAAA 180
QY 181 CGCGGAAAGAGTGAAGCGCGATGCGGAGCTGAATTAACATTCCTCCACCGCTGGGCAC 240
Db 181 CGCGGAAAGAGTGAAGCGCGATGCGGAGCTGAATTAACATTCCTCCACCGCTGGGCAC 240
QY 241 AACAACTGCGGGCAACACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGCAACACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTCGCGCGCATTAATAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTCGCGCGCATTAATAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGATGTAGTAAACGAAAGCGCGCTGAAAGCTGTAAAGCGCGGTGCCAATC 420
Db 361 TGGTGTGTGATGTAGTAAACGAAAGCGCGCTGAAAGCTGTAAAGCGCGGTGCCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCGGCGCTTATTTCTGATGTCTCTGACCA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCGGCGCTTATTTCTGATGTCTCTGACCA 540
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACCGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACCGGACTGGCGGTGGAGCATC 600
```

Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
Qy 601 TGGTCGCAATTGGGTCAACGAAATCGCGCTGTGTAGCGGCCCATTTAAAGTCTCTGCTCGG 660
Db 601 TGGTCGCAATTGGGTCAACGAAATCGCGCTGTGTAGCGGCCCATTTAAAGTCTCTGCTCGG 660
Qy 661 CGCGTCTGGCTGTGGCTGGCTGGCTGAATAATATCTCATCTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTGTGGCTGGCTGGCTGAATAATATCTCATCTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGCACTGAGTGGCATGTCGGTTCCTCAACGATGGCGCTGGCGCGAA 840
Db 721 CGGAACGGGAAGCGCACTGAGTGGCATGTCGGTTCCTCAACGATGGCGCTGGCGCGAA 840
Qy 781 ATGAGGGCATCTGTTCCCACTGGGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCTGTTCCCACTGGGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCGAA 840
Qy 841 TGGCGGCCATTACCGAGTCCGGCTGGCGTGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTACCGAGTCCGGCTGGCGTGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGGCGCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGGCGCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGCTGGCGGATTCATTAATGAGCTGSCACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTGGCGGATTCATTAATGAGCTGSCACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTAAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTAAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCAACCGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCAACCGTT 1380
Qy 1381 CTGGCAATATTTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGGCAATATTTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Qy 1441 ATTGTAGCGGATTAACAATTTCAACAGGAACAGCCAGTCCGTTTAGGTGTTTTTCACGA 1500
Db 1441 ATTGTAGCGGATTAACAATTTCAACAGGAACAGCCAGTCCGTTTAGGTGTTTTTCACGA 1500
Qy 1501 GCATTCACCAACAGGACCATAGATTATGAAACTGAAGAGTAACCTGGTAAATCTGG 1560
Db 1501 GCATTCACCAACAGGACCATAGATATGAAATTCGAAGAGTAACCTGGTAAATCTGG 1560
Qy 1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAACTCGGTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAACTCGGTAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680

RESULT 12

AD023639

ID AD023639 standard; DNA; 7370 BP.

Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGCAACAAAGGTTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGCAACAAAGGTTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACTTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGTTTACC CGATCGCTGTT 1860
Db 1801 CCGTTTACTTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGTTTACC CGATCGCTGTT 1860
Qy 1861 GAACGGTTATCGCTGATTTTATACAAAGATCTGCTGCCAACC CGCCAAAACCTGGGAA 1920
Db 1861 GAACGGTTATCGCTGATTTTATACAAAGATCTGCTGCCAACC CGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGGCAAGTACGACATTTAAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTTAAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCGCAAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Qy 2221 GCATGTGCCAAACATCAACACAGCAAGTGAATTTATGTTGTTAAACGCTGCTGCGACCTTC 2280
Db 2221 GCATGTGCCAAACATCAACACAGCAAGTGAATTTATGTTGTTAAACGCTGCTGCGACCTTC 2280
Qy 2281 AAGGTCACCATTCACAAACCGTTCTGTTGGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCATTCACAAACCGTTCTGTTGGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGCAAAAGAGTTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGCAAAAGAGTTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTCGCCGACCATGATGGAAGAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTCGCCGACCATGATGGAAGAACCGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGTGGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGTGGCTGCTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGCTGAGTCTGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGCTGAGTCTGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACCAACAAACAATAACCAATAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAATAACCAATAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700

```

XX AC ADO23639;
XX AC
XX DT 01-JUL-2004 (first entry)
XX DE
XX DE DNA encoding MBP-ToxoP30MIX1 fusion protein.
XX KW P30 antigen; Toxo30del13C; Toxo30del2C; ToxoP30 MIX1;
XX KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
XX KW Toxoplasma gondii; ds; gene.
XX OS
XX OS Toxoplasma gondii.
XX OS Synthetic.
XX PW US2004067239-A1.
XX PD
XX PD 08-APR-2004.
XX PF
XX PF 02-OCT-2002; 2002US-00263153.
XX PR
XX PR 02-OCT-2002; 2002US-00263153.
XX PA (MAIN/) MAINE G T.
XX PA (PATE/) PATEL C B.
XX PA (GINS/) GINSBURG S R.
XX PA (BLIE/) BLIESE T R.
XX PA
XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX XX
XX XX WPI; 2004-304563/28.
XX DR P-PSDB; ADO23640.
XX XX
XX PT Novel purified polypeptide having sequence identity to amino acid
XX PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
XX PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
XX PT antibodies to Toxoplasma gondii.
XX XX
XX PS Example 5; Fig 31; 114pp; English.
XX XX
XX CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
XX CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
XX CC amino acids added to the C-terminus of the amino acid sequence of
XX CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
XX CC which at least one of the five C-terminal cysteine amino acids of the
XX CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
XX CC with alanine, or comprising the amino acid sequence chosen from MBP-
XX CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
XX CC fusion proteins. (I) is useful for detecting the presence of IgM
XX CC antibodies to Toxoplasma gondii in a test sample, which involves
XX CC contacting the test sample suspected of containing the IgM antibodies
XX CC with a composition comprising (I) and detecting the presence of (I)/IgM
XX CC antibody complexes. The present sequence represents DNA encoding a MBP-
XX CC ToxoP30 fusion protein of the invention.
XX XX
XX SQ Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;

Query Match      81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGCGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGCGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTACAGGTGTGAATGTGAACAGTAACGTTATACGATGTCGCGAGATATCGCG 120
DB 61 GTCAATTACAGGTGTGAATGTGAACAGTAACGTTATACGATGTCGCGAGATATCGCG 120
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGGGTGTGAACAGGCCAGCCACGCTTTCCTGGGAAAA 180
DB 121 GTGTCTCTTATCAGACCGGTTTCCCGGGTGTGAACAGGCCAGCCACGCTTTCCTGGGAAAA 180
QY 181 CCGGGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAAATTACATTTCCCAACCGCTGGGCAC 240

```

```

Db 181 CCGGGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAAATTACATTTCCCAACCGCTGGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCCAATTTGTCGGGGGATTAATCTCGCGCGGATCAATCTGGGTGCGAGG 360
Db 301 ACGCGCGTCCCAATTTGTCGGCGGATTAATCTCGCGCGGATCAATCTGGGTGCGAGG 360
QY 361 TGGTGGTGTGATGATGAAGCGCGCTGGAAGCGCTGTAAGCGCGTGAAGCGCGTGCACAAATC 420
Db 361 TGGTGGTGTGATGATGAAGCGCGCTGGAAGCGCTGTAAGCGCGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCA 540
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTGGCATTTGGGTCAACGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCG 660
Db 601 TGGTGGCATTTGGGTCAACGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCG 660
QY 661 CGCGTCTCGCTCTGGCTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTCGCTCTGGCTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGGCAA 840
QY 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGGAAGACGCTCATGTTATATCCCGCGGTTAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGGAAGACGCTCATGTTATATCCCGCGGTTAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAACACGAGCTGGAACCGCTTGTGCAACTCTCTCAGGGCCAGGGGGTGA 1020
Db 961 GCCTGCTGGGGCAACACGAGCTGGAACCGCTTGTGCAACTCTCTCAGGGCCAGGGGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCACTGTTGTAAGAAAGAAAACCACTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGCTCACTGTTGTAAGAAAGAAAACCACTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCGACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGAACAGCTTATCATCGACTGCGGTCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGAACAGCTTATCATCGACTGCGGTCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAGCCATCGGAAGCTGTGTTATGCTGTGAGGTGCTGAATCACTGCATTAATTCG 1320

```

Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAGTGTGTAATCACTGCATATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCCTGGATATGTTTTTTCGCCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCCTGGATATGTTTTTTCGCCGACATCATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTGTGAATATCAATCATCGGCTCGTATATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTGTGAATATCAATCATCGGCTCGTATATGTGTGA 1440
QY 1441 ATTGTAGCGGATACAAATTTTTCACACAGGAACAGCCAGTCCGTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTTCACACAGGAACAGCCAGTCCGTTAGTGTTCACGA 1500
QY 1501 GCATTTCCAAACAAGGACCATAGATATTAAGAACTGAAGAGGTAAATCTGTAATCTGG 1560
Db 1501 GCATTTCCAAACAAGGACCATAGATATTAAGAACTGAAGAGGTAAATCTGTAATCTGG 1560
QY 1561 ATTAACGGGATAAAGGCTATACCGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATACCGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATCAATCAATCGGTAAGTAAAGAAATTCGAGAAAGAT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATCAATCAATCGGTAAGTAAAGAAATTCGAGAAAGAT 1680
QY 1681 GCGGCAACTGCGATGCGCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGCGATGCGCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAATATCAACCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAATATCAACCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACTCTGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACTCTGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCGATCGCTGTT 1860
QY 1861 GAAGGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCGCCCAAAACCTGGGAA 1920
QY 1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGCGCGAAACG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGCGCGAAACG 2100
QY 2101 GGTCTGACCTTCCTGGTTGACCTGATTAAGAAACAAACACATGATGAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCCTGGTTGACCTGATTAAGAAACAAACACATGATGAGACCGGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTATAAGGCGAAACAGCGATGACCATCAACCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTATAAGGCGAAACAGCGATGACCATCAACCGCCGCTGG 2220
QY 2221 GCATGTTCCACATCGACACAGCAAGTGAATATGTTGATACCGTTATCGCGACCTTC 2280
Db 2221 GCATGTTCCACATCGACACAGCAAGTGAATATGTTGATACCGTTATCGCGACCTTC 2280
QY 2281 AAGGTTCAACATCAACACCGTTGCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCACT 2340
Db 2281 AAGGTTCAACATCAACACCGTTGCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCACT 2340
QY 2341 CCGAACCAAGAGCTGGCAAAAGATTCCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
Db 2341 CCGAACCAAGAGCTGGCAAAAGATTCCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400

QY 2401 GAACGGTTAATAAGACAAACCCCTGGTGCCTAGCGTGAAGTCTTACGAGAAAG 2460
Db 2401 GAACGGTTAATAAGACAAACCCCTGGTGCCTAGCGTGAAGTCTTACGAGAAAG 2460
QY 2461 TTGCGGAAAGATCCACGTATTTCGCCGACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTATTTCGCCGACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGGTATGCCGTGCTACTCGCGTGATCAAGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGGTATGCCGTGCTACTCGCGTGATCAAGCC 2580
QY 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGTCTG 2640
Db 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGTCTG 2640
QY 2641 AACAAACAACAATAAATAAACAACCTCGGGATCGAGGAGAGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAATAAACAACCTCGGGATCGAGGAGAGATTCAGAAATTC 2700

RESULT 13
ADO23649

ID ADO23649 standard; DNA; 7370 BP.

XX ADO23649;

XX AC ADO23649;

XX DT 01-JUL-2004 (first entry)

XX DE DNA encoding MBP-Toxop30MIX5 fusion protein.

XX KW P30 antigen; Toxo30del13C; Toxo30del12C; Toxop30 MIX1;

XX KM MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;

XX KW Toxoplasma gondii; ds; gene.

XX OS Toxoplasma gondii.

XX OS Synthetic.

XX PN US2004067239-A1.

XX PD 08-APR-2004.

XX PF 02-OCT-2002; 2002US-00263153.

XX PR 02-OCT-2002; 2002US-00263153.

XX PA (MAIN/) MAINE G T.

XX PA (PATE/) PATEL C B.

XX PA (GINS/) GINSBURG S R.

XX PA (BLIE/) BLIESE T R.

XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;

XX WI MPI; 2004-304563/28.

XX DR P-PSDB; ADO23650.

XX XX Novel purified polypeptide having sequence identity to amino acid

XX PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,

XX PT Toxo30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG

XX PT antibodies to Toxoplasma gondii.

XX PS Example 5; Fig 35; 114pp; English.

XX CC The invention relates to a purified P30 antigen (I) chosen from 3 fully

XX CC defined Toxo30del13C, Toxo30del12C and Toxop30 MIX1 sequences, having 1-6

XX CC amino acids added to the C-terminus of the amino acid sequence of

XX CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in

XX CC which at least one of the five C-terminal cysteine amino acids of the

XX CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted

XX CC with alanine, or comprising the amino acid sequence chosen from MBP-

XX CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1

XX CC fusion proteins. (I) is useful for detecting the presence of IgM

CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ

Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;	
Query Match	81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 2694; Conservative	0; Mismatches 6; Indels 0; Gaps 0;
QY	1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATATAGCGCCCGGAAGAGA 60
Db	1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATATAGCGCCCGGAAGAGA 60
QY	61 GTCGAATTCAGGGTGGTGAATGTGAACACGATTAACGATATACGATGTCGACAGATATGCCG 120
Db	61 GTCGAATTCAGGGTGGTGAATGTGAACACGATTAACGATATACGATGTCGACAGATATGCCG 120
QY	121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAAA 180
Db	121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAAA 180
QY	181 CGCGGAAAAAGTGAAGCGGGATGCGGAGCTGAATTAATTCCTCCACCGCTGGCAC 240
Db	181 CGCGGAAAAAGTGAAGCGGGATGCGGAGCTGAATTAATTCCTCCACCGCTGGCAC 240
QY	241 AACAACTCGCGGGAACAGATGCTTGTGATTTGGCGTTTGCACCTCAGTCTGGGCCCTGC 300
Db	241 AACAACTCGCGGGAACAGATGCTTGTGATTTGGCGTTTGCACCTCAGTCTGGGCCCTGC 300
QY	301 ACGCGCGTCGCAAAATGTCGCGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db	301 ACGCGCGTCGCAAAATGTCGCGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY	361 TGGTGGTTCGATGTAGAACGAAGCGCGTTCGAAGCCTGTAAAGCGCGTGCACAATC 420
Db	361 TGGTGGTTCGATGTAGAACGAAGCGCGTTCGAAGCCTGTAAAGCGCGTGCACAATC 420
QY	421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACACGAGATGCCA 480
Db	421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACACGAGATGCCA 480
QY	481 TTGCTGTGGAAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCA 540
Db	481 TTGCTGTGGAAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCA 540
QY	541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGGGACTGGGCGTGGAGCATC 600
Db	541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGGGACTGGGCGTGGAGCATC 600
QY	601 TGGTCGCAATGGGTCCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db	601 TGGTCGCAATGGGTCCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
QY	661 CGCGTCTCGCTGCGCTGGCGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db	661 CGCGTCTCGCTGCGCTGGCGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY	721 CGGAAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db	721 CGGAAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY	781 ATGAGGGCATGTTCCCACTCGGATGCTGGTTCGCAACGATCAGATGGCGTGGCGGCA 840
Db	781 ATGAGGGCATGTTCCCACTCGGATGCTGGTTCGCAACGATCAGATGGCGTGGCGGCA 840
QY	841 TGC CGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db	841 TGC CGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY	901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACGAGATTTC 960

Db	901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACGAGATTTC 960
QY	961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db	961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
QY	1021 AGGCAATCAGCTGTTGCCGCTCTCACTGTTGAAGAAAGAAACCAACCTGGGGCCCAATA 1080
Db	1021 AGGCAATCAGCTGTTGCCGCTCTCACTGTTGAAGAAAGAAACCAACCTGGGGCCCAATA 1080
QY	1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCAGACAGAGTTT 1140
Db	1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCAGACAGAGTTT 1140
QY	1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db	1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY	1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCG 1260
Db	1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCG 1260
QY	1261 TCAGCAGCCATCGAAAGCTGTGTATGCTGTGAGCTGCTAAATCACTGCATAAATTCG 1320
Db	1261 TCAGCAGCCATCGAAAGCTGTGTATGCTGTGAGCTGCTAAATCACTGCATAAATTCG 1320
QY	1321 TGTGCTCAAGCGCACCTCCCGTCTTGGATTAATGTTTTTTCGCGCGACATCAACGGTT 1380
Db	1321 TGTGCTCAAGCGCACCTCCCGTCTTGGATTAATGTTTTTTCGCGCGACATCAACGGTT 1380
QY	1381 CTGGCAAAATCTTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Db	1381 CTGGCAAAATCTTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
QY	1441 ATTGTAGCGGATCAACAAATTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAG 1500
Db	1441 ATTGTAGCGGATCAACAAATTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAG 1500
QY	1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAAAGGTAAATCTGTTATCTGG 1560
Db	1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAAAGGTAAATCTGTTATCTGG 1560
QY	1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAACTCGGTGAAGAAATTCGAGAAAG 1620
Db	1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAACTCGGTGAAGAAATTCGAGAAAG 1620
QY	1621 ACCGGAAATTAAGTCAACCGTTGACATCCCGATAACTGGAAGAGAAATTCACAGGTT 1680
Db	1621 ACCGGAAATTAAGTCAACCGTTGACATCCCGATAACTGGAAGAGAAATTCACAGGTT 1680
QY	1681 CGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db	1681 CGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
QY	1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAAGCTGAT 1800
Db	1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAAGCTGAT 1800
QY	1801 CCGTTTACCTGGGATGGCTAGCTTACAAACCGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db	1801 CCGTTTACCTGGGATGGCTAGCTTACAAACCGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY	1861 GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCCAAAACCTGGGAA 1920
Db	1861 GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCCAAAACCTGGGAA 1920
QY	1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db	1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY	1981 CTGCAGAACCGTACTTCACTTGGCGCTGATTGCTGCTGACGGGGTTATCGGTTCAAG 2040

```
Db 1981 CTGCAAGAACCGTACTTCACTACCTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAAACAACACATGAATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAAACAACACATGAATGACAGACCGGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACAGCAGTACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACAGCAGTACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACACATCCAAACCGTTCTGGCGTGTGAGCGAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTCACACATCCAAACCGTTCTGGCGTGTGAGCGAGGTATTAAACGCCGCCAGT 2340
QY 2341 CGAACAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CGAACAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAAGTCTTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGTTATTGCGGCCACCATGGAAGAGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATTGCGGCCACCATGGAAGAGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATACAAATAACAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATACAAATAACAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

RESULT 14
ID ADO23644
XX ADO23644 standard; DNA; 7370 BP.
AC ADO23644;
XX
DT 01-JUL-2004 (first entry)
DE DNA encoding MBP-Toxop30MIX3 fusion protein.
XX
KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
(PATN/) MAINE G T.
```

```
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23645.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of Igm or Igg
PT antibodies to Toxoplasma gondii.
XX
PS Example 5; Fig 33; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of Igm
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the Igm antibodies
CC with a composition comprising (I) and detecting the presence of (I)/Igm
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGCGCCCGAAGAGA 60
QY 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAACTATACATGTCGAGAGTATGCGCG 120
Db 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAACTATACATGTCGAGAGTATGCGCG 120
QY 121 GTGTCTTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTATGTCGAGAGTATGCGCG 180
Db 121 GTGTCTTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTATGTCGAGAGTATGCGCG 180
QY 181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAAAAGTGTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAAGTGTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATTAATCTCGCGCGGATTAATCTCGCGCGG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATTAATCTCGCGCGGATTAATCTCGCGCGG 360
QY 361 TGGTGGTGTGATGGTGAACGAAGCGCGTGAAGCGCTGTAAAGCGCGGTGCAACATC 420
Db 361 TGGTGGTGTGATGGTGAACGAAGCGCGTGAAGCGCTGTAAAGCGCGGTGCAACATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTTAATTTTCGGCGGTTATTTCTGATGTCCTGACCGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTTAATTTTCGGCGGTTATTTCTGATGTCCTGACCGA 540
```

QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACCGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACCGGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCATTGGGTACACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTACACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGGCTGGCTGGCGTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTGGCTGGCGTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCACCTGCGATGCTGTGTGCCAAACGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCGTTCACCTGCGATGCTGTGTGCCAAACGATCAGATGGCGCTGGCGCGAA 840
QY 841 TGGCGCCATTACCAAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCCATTACCAAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACGGAAGACGCTCATGTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACGGAAGACGCTCATGTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGGTGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGGTGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCCCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGTAGTCACTCATTTAG 1140
Db 1081 CGCAAAACCCCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGTAGTCACTCATTTAG 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTGTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTGTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGAACGCTTATCATGCACTGCACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGAACGCTTATCATGCACTGCACGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAGGCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGCAGGCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTTCGCGCAGCATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTTCGCGCAGCATCAATACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATACAATTTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTCAGA 1500
Db 1441 ATTGTAGCGGATACAATTTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTCAGA 1500
QY 1501 GCACCTTCCACCAAGGACCATAGCATATGAAAATCGAAGAGGTAAACTGCTGATCTCG 1560
Db 1501 GCACCTTCCACCAAGGACCATAGCATATGAAAATCGAAGAGGTAAACTGCTGATCTCG 1560
QY 1561 ATTTAAACGGGATAAAGGCTTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAAACGGGATAAAGGCTTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTTCCACAGGTT 1680

Db 1621 ACCGGAAATTAAAGTCACCGTTGAGCATCCGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGGCAACTGCGGATGGCCCTTGACATATCTTCTTGGGCACACGACCGCTTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGGATGGCCCTTGACATATCTTCTTGGGCACACGACCGCTTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGGACAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGGACAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCGTACGTTTACAACCGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGGCGTACGTTTACAACCGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAACGGAAAGSTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAACGGAAAGSTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTTAAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTTAAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAAACATGAAATGCAAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAAACATGAAATGCAAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTAAAGCGCAACAGCGATGACCATCAAGCGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTTAAAGCGCAACAGCGATGACCATCAAGCGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACACGACCAAGTGAATTAATGTTGTAACCGTACTGCGCACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGACCAAGTGAATTAATGTTGTAACCGTACTGCGCACCTTC 2280
QY 2281 AAGGTCACCAATCAACACCGTTGCGTGTGAGCGCAGGATTAATGAGCCGCCAGT 2340
Db 2281 AAGGTCACCAATCAACACCGTTGCGTGTGAGCGCAGGATTAATGAGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTTAAATAAGACAAACCCGCTGGGTGCGGTAGCGCTGAACTTTTACGAGGAAG 2460
Db 2401 GAAGCGTTTAAATAAGACAAACCCGCTGGGTGCGGTAGCGCTGAACTTTTACGAGGAAG 2460
QY 2461 TTGGCAAAAGATTCACGTTATTTGGCGCCACCATGGAAGAAACGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCAAAAGATTCACGTTATTTGGCGCCACATATGGAAGAAACGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGGCGACACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGGCGACACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAATAACAACTTCGCGGATCGAGGAAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAATAACAACTTCGCGGATCGAGGAAAGGATTTTCAGAAATTC 2700

AD023594
ID ADO23594 standard; DNA; 7403 BP.
XX
AC ADO23594;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30del12(52-311aa) fusion protein.
XX
KW P30 antigen; Tox30del13C; Tox30del12C; Toxop30 MIX1;
KW MBP-Tox30del13C(52-300aa); MBP-Tox30del14C(52-294aa); MBP-Tox30MIX1;
KW Toxoplasma gondii; da; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
DR WPI; 2004-304563/28.
DR P-PSDB; ADO23595.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Tox30del13C,
PT Tox30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
PS Example 2; Fig 11; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Tox30del13C, Tox30del12C and Toxop30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Tox30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Tox30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Tox30del13C(52-300aa), MBP-Tox30del14C(52-294aa) and MBP-Tox30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
SQ Sequence 7403 BP; 1907 A; 1863 C; 1961 G; 1672 T; 0 U; 0 Other;

Query Match
Best Local Similarity 81.5%; Score 2690.4; DB 12; Length 7403;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGATGTGGAATGTGAACACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGTGTGAATGTGAACACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTACAGGCGGTGTGAATGTGAACACCTTTCGCGGTATGGCATGATAGCGCGATATGCCG 120
DB 61 GTCAATTACAGGCGGTGTGAATGTGAACACCTTTCGCGGTATGGCATGATAGCGCGATATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGAACACCGCGGTGTGAACACCGCGGTGTGAACACCG 180
DB 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGAACACCGCGGTGTGAACACCGCGGTGTGAACACCG 180

QY 181 CGCGGGAAGTGAAGCGCGATGGCGGAGCTGAATTAATCATTTCCCAACCGCGTGGCAC 240
DB 181 CGCGGGAAGTGAAGCGCGGATGGCGGAGCTGAATTAATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTGTGTGTGATGGCGGTGGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTGTGTGTGATGGCGGTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCCAATTTGTTCGGCGGATTAATCTCCGGCGGATCAATCTGGGTGCCAGCG 360
DB 301 ACGCGCGTCCCAATTTGTTCGGCGGATTAATCTCCGGCGGATCAATCTGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGTAGAACGAAGCGCGTGAAGCTCTGAAGCGCGGTGCACATC 420
DB 361 TGGTGGTTCGATGTAGAACGAAGCGCGTGAAGCTCTGAAGCGCGGTGCACATC 420
QY 421 TTCTCGCGCACCGCTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480
DB 421 TTCTCGCGCACCGCTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGAAGTGCCTGCATTAATGTTCCGCGTATTTCTTGATGTCTTGACAGAGA 540
DB 481 TTGCTGTGAAGTGCCTGCATTAATGTTCCGCGTATTTCTTGATGTCTTGACAGAGA 540
QY 541 CACCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATC 600
DB 541 CACCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATC 600
QY 601 TGGTTCGCAATTCAGCAATCGCTGTAGCGGCGCATTAAGTCTGTCTCGG 660
DB 601 TGGTTCGCAATTCAGCAATCGCTGTAGCGGCGCATTAAGTCTGTCTCGG 660
QY 661 CGGCTGTGGCTTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
DB 661 CGGCTGTGGCTTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
QY 721 CGGAACGGGAAGCGACTCGAGTGCATGTCGGTTCGATTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGACTCGAGTGCATGTCGGTTCGATTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCAATCGTTCCTCCACTGGATGCTGGTTCGCAACGATCAGATGGCTGGCGCAA 840
DB 781 ATGAGGCAATCGTTCCTCCACTGGATGCTGGTTCGCAACGATCAGATGGCTGGCGCAA 840
QY 841 TCGCGCCATTACCGAGTCCGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGATACG 900
DB 841 TCGCGCCATTACCGAGTCCGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGCACCGTTCGCAACTCTCTCAGGCGCAGCGGTGA 1020
DB 961 GCCTGCTGGGCAAAACAGCGTGCACCGTTCGCAACTCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTGGCCGATTCATTAATGACAGTGGCAGCAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTGGCCGATTCATTAATGACAGTGGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGCAGCAATTAATGATGATGCTCACTATTAG 1200
DB 1141 CCCGACTGGAAAGCGGCGAGTGCAGCAATTAATGATGATGCTCACTATTAG 1200
QY 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260

QY 1261 TCAGCGAGCCATCGAAGCTGTGGTATGGCTGTGCGAGGTGCTAATCACTGCTAATTCG 1320
Db 1261 TCAGCGAGCCATCGAAGCTGTGGTATGGCTGTGCGAGGTGCTAATCACTGCTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGATATGTTTTTGGCGCGACATCATACGGTT 1380
QY 1381 CTGGCAATATTTGAAATGAGCTGTGACATTAATCATCGGCTCGTATATGTTGTGA 1440
Db 1381 CTGGCAATATTTGAAATGAGCTGTGACATTAATCATCGGCTCGTATATGTTGTGA 1440
QY 1441 ATTGTGAGCGGTAACAAATTTACACAGGAAACAGCCAGTCCGTTAGGTTGTTTCAG 1500
Db 1441 ATTGTGAGCGGTAACAAATTTACACAGGAAACAGCCAGTCCGTTAGGTTGTTTCAG 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTTATGAAACTGAAAGAGTAAACTGCTAATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTTATGAAATTCGAAGAGGTAAACTGGTAACT 1560
QY 1561 ATTTAAACGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAA 1620
Db 1561 ATTTAAACGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAA 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTGGGCACACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTGGGCACACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTAGCTTACACCGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTACACCGCAAGCTGATGCTTACCCGATCGCTGT 1860
QY 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
QY 1921 GAGATCCGCGCTGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGGCTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGGCTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 2160
Db 2101 GGTCTGACCTTCTGCTGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 2160
QY 2161 TCCATCGCAAGAGTCCCTTAAATGAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAAGAGTCCCTTAAATGAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACGCAAGTAAATGATGTTAAACGCTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACGCAAGTAAATGATGTTAAACGCTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400

Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCCGCTGGTGCCGTAGCGCTGAAGTCTTACGAGGAAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCCGCTGGTGCCGTAGCGCTGAAGTCTTACGAGGAAAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTTGGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTTGGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACATTAACATTAACAAACCTCCGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAAACATTAACATTAACAAACCTCCGGATCGAGGGAAGGATTTCAAGATTC 2700

Search completed: August 1, 2005, 04:28:59
Job time : 1101.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-18

Perfect score: 3300

Sequence: 1 ccgacacacgaatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2700	81.8	7475	2 US-08-971-036-1	Sequence 1, Appli
2	2700	81.8	7475	3 US-09-096-570-1	Sequence 1, Appli
3	2700	81.8	7475	3 US-09-265-617B-1	Sequence 1, Appli
c 4	1417.6	43.0	5926	3 US-09-027-169-3	Sequence 3, Appli
5	1362	41.3	5201	4 US-09-640-882-2	Sequence 2, Appli
6	1362	41.3	5201	4 US-09-640-882-3	Sequence 3, Appli
7	1204	36.5	4557	4 US-08-778-717-5	Sequence 5, Appli
8	1200.8	36.4	3832	1 US-08-148-675A-2	Sequence 2, Appli
c 9	1199.2	36.3	5248	3 US-08-487-283A-18	Sequence 18, Appli
c 10	1199.2	36.3	5248	5 PCT-US96-05611A-21	Sequence 21, Appli
c 11	1199.2	36.3	5312	4 US-10-263-103-35	Sequence 35, Appli
c 12	1199.2	36.3	5443	2 US-08-929-967-1	Sequence 1, Appli
c 13	1199.2	36.3	5502	4 US-09-702-705-785	Sequence 785, App
c 14	1199.2	36.3	5502	4 US-09-736-457-785	Sequence 785, App
c 15	1199.2	36.3	5502	4 US-09-614-124B-785	Sequence 785, App
c 16	1199.2	36.3	5502	4 US-09-671-325-785	Sequence 785, App
c 17	1199.2	36.3	5502	4 US-09-589-184-785	Sequence 785, App
c 18	1199.2	36.3	5502	4 US-09-658-824-785	Sequence 785, App
c 19	1199.2	36.3	5616	2 US-08-929-967-3	Sequence 3, Appli
c 20	1199.2	36.3	5873	4 US-09-695-437A-62	Sequence 62, Appli
c 21	1199.2	36.3	6353	4 US-09-702-705-784	Sequence 784, App
c 22	1199.2	36.3	6353	4 US-09-736-457-784	Sequence 784, App
c 23	1199.2	36.3	6353	4 US-09-614-124B-784	Sequence 784, App
c 24	1199.2	36.3	6353	4 US-09-671-325-784	Sequence 784, App
c 25	1199.2	36.3	6353	4 US-09-589-184-784	Sequence 784, App
c 26	1199.2	36.3	6353	4 US-09-658-824-784	Sequence 784, App
c 27	1199.2	36.3	6361	3 US-08-646-538-7	Sequence 7, Appli

c 28	1199.2	36.3	6361	3 US-09-503-222-7	Sequence 7, Appli
c 29	1199.2	36.3	6727	2 US-08-125-462-2	Sequence 2, Appli
c 30	1199.2	36.3	6727	2 US-08-891-848-2	Sequence 2, Appli
c 31	1199.2	36.3	6799	2 US-08-125-462-5	Sequence 5, Appli
c 32	1199.2	36.3	6799	2 US-08-891-848-5	Sequence 5, Appli
c 33	1199.2	36.3	7659	3 US-09-128-314-4	Sequence 4, Appli
c 34	1199.2	36.3	7676	3 US-09-056-556-213	Sequence 213, App
c 35	1199.2	36.3	7676	3 US-08-072-596-208	Sequence 208, App
c 36	1199.2	36.3	7676	4 US-09-072-967-213	Sequence 213, App
c 37	1199.2	36.3	7676	4 US-09-287-849-9	Sequence 9, Appli
c 38	1199.2	36.3	8031	3 US-09-643-597-254	Sequence 254, App
c 39	1199.2	36.3	8031	4 US-09-480-884A-254	Sequence 254, App
c 40	1199.2	36.3	8031	4 US-09-542-615A-254	Sequence 254, App
c 41	1199.2	36.3	8031	4 US-08-606-421B-254	Sequence 254, App
c 42	1199.2	36.3	8031	4 US-09-476-496A-254	Sequence 254, App
c 43	1199.2	36.3	8031	4 US-09-630-940B-254	Sequence 254, App
c 44	1199.2	36.3	8157	3 US-09-128-314-3	Sequence 3, Appli
c 45	1199.2	36.3	8501	3 US-08-793-900-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8%; Score 2700; DB 2; Length 7475;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CGGACACCATCAATGGTGCAGAACTTTCGGCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
Db 1 CGGACACCATCAATGGTGCAGAACTTTCGGCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGGTGTAACACAGCCAGCCACGCTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGGTGTAACACAGCCAGCCACGCTTCTCGGAAAA 180
QY 181 CCGCGGAAAAAGTGAAGCGGGATGGCGGAGCTGAATTAACATCCCAACCGCTGGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGGGATGGCGGAGCTGAATTAACATCCCAACCGCTGGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTGTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTGTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTTGTCCGGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGC 360
Db 301 ACGCGCGTTCGCAAAATTTGTCCGGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGC 360
QY 361 TGGTGTGTCGATGGTGAACGAAGCGCGTGAAGCGCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGTGTCGATGGTGAACGAAGCGCGTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGGGTGAGTGGGTGATCAATACTATCCGCTGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACGGGTGAGTGGGTGATCAATACTATCCGCTGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTGACACAGA 540
Db 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTGACACAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGACTGGGCGTGGAGATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGACTGGGCGTGGAGATC 600
QY 601 TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCCTCTCGG 660
Db 601 TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCCTCTCGG 660
QY 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTAGCGGGCCCAATTAAGTTCCTCTCGG 840
Db 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTAGCGGGCCCAATTAAGTTCCTCTCGG 840
QY 841 TGGCGGCAATTAACGAGTTCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCAATTAACGAGTTCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGTCTATGTTATATCCGCGGTTTAAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGTCTATGTTATATCCGCGGTTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGTGSGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db 961 GCCTGTGSGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGCAATACGCTGTTCGCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATACGCTGTTCGCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
```

```
Db 1021 AGGCAATACGCTGTTCGCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGAAAAACCGCTCTTCCCGCGCGGTGGCGGATTCATTAATGCACTGGCAGCAGCAGGTTT 1140
Db 1081 CGAAAAACCGCTCTTCCCGCGCGGTGGCGGATTCATTAATGCACTGGCAGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
QY 1201 GCACAAATCTCATGTTTTGACAGCTTATCATCGACTGACCGGTGCACCAATGCTTCTGCGG 1260
Db 1201 GCACAAATCTCATGTTTTGACAGCTTATCATCGACTGACCGGTGCACCAATGCTTCTGCGG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGATAGGCTGTGTCAGGTCGTAAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGATAGGCTGTGTCAGGTCGTAAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGATAAATGTTTTTTCGCCCGGACATCAACCGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGATAAATGTTTTTTCGCCCGGACATCAACCGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
QY 1441 ATTGTAGCGCGATAAACATTTTCACAGGAAACAGCCAGTCCGTTTATAGTGTTCACGA 1500
Db 1441 ATTGTAGCGCGATAAACATTTTCACAGGAAACAGCCAGTCCGTTTATAGTGTTCACGA 1500
QY 1501 GCACCTTCCACCAACAAAGGACCATAGATTTATGAAACTGGAAGAGTAACTGTTAACTCGG 1560
Db 1501 GCACCTTCCACCAACAAAGGACCATAGATTTATGAAACTGGAAGAGTAACTGTTAACTCGG 1560
QY 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGTGAAGTGGTAAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGTGAAGTGGTAAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACAGCCGCTTTGGTGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACAGCCGCTTTGGTGCTAC 1740
QY 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCCCGGAACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCCCGGAACAAAGCGTTCCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTTACCTGGGATCCGTTACGTTAAACCGCAAGCTGATTCGTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTTACCTGGGATCCGTTACGTTAAACCGCAAGCTGATTCGTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTCGGAA 1920
QY 1921 GAGATCCCGCGCTGATTAAGAACTGAAGAGTGAAGAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGATTAAGAACTGAAGAGTGAAGAGCGCGCTGATGTTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTCGTTGACGGGGTTTATGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTCGTTGACGGGGTTTATGGTTCAAG 2040
QY 2041 TATGAAGAACCGGAGTACGACATTAAGAGACGTGGCGTGGATAAACGCTGGCGGAAGCG 2100
Db 2041 TATGAAGAACCGGAGTACGACATTAAGAGACGTGGCGTGGATAAACGCTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAGAAACCAACATCAATGAATGCAGACCCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAGAAACCAACATCAATGAATGCAGACCCGATTAC 2160
```

Qy	2161	TCCATCCGAGAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGCTGG	2220
Db	2161	TCCATCCGAGAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGCTGG	2220
Qy	2221	GCATGGTCCAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGGTCCAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGGTCAACGATCCAAACCGTTGCTGGGGTGTCTGAGCGCAGGTATTAACGGCCGCCAGT	2340
Db	2281	AAGGGTCAACGATCCAAACCGTTGCTGGGGTGTCTGAGCGCAGGTATTAACGGCCGCCAGT	2340
Qy	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCCTG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCCTG	2400
Qy	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGCGGAAAGATCCACGCTATTGCCGCCACCATGGAAAAACGCCCAAGAGGTGAATCATG	2520
Db	2461	TTGCGGAAAGATCCACGCTATTGCCGCCACCATGGAAAAACGCCCAAGAGGTGAATCATG	2520
Qy	2521	CCGAAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCTACTCGGCTGATCAACGCC	2580
Db	2521	CCGAAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCTACTCGGCTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCGACACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCGACACTAATTCGAGCTCG	2640
Qy	2641	AACAAACAACAAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGATTC	2700
Db	2641	AACAAACAACAAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGATTC	2700

RESULT 2

```

US-09-056-570-1
; Sequence 1, Application US/09096570
; Patent No. 6018020
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.2.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/096,570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522

```

Qy	841	TGCGCGCATTACAGAGTCGGGCTCGCGTTGGTGGCGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGGGCATTTACGAGTCCGGGCTCGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTTATATCCGCGCGTTAACACCACCATCAAAACGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTTATATCCGCGCGTTAACACCACCATCAAAACGATTTTC	960
Qy	961	GCTTGCTGGGGCAACACAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCTTGCTGGGGCAACACAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTGTCGCCGCTCACTGTGTGAAGGAAAGCAACACCTGCGGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTGTCGCCGCTCACTGTGTGAAGGAAAGCAACACCTGCGGCCCAATA	1080
Qy	1081	CGCAAAACCGCCTCTCCCGCGCGTGTGGCCGATTCATTAATGCAGCTGGCAGCAGGTTT	1140
Db	1081	CGCAAAACCGCCTCTCCCGCGCGTGTGGCCGATTCATTAATGCAGCTGGCAGCAGGTTT	1140
Qy	1141	CCCGACTGGAAGCGGCGAGTGACGGCAACGCCAATTAATGTGAGTTAGTCACTCATTAG	1200
Db	1141	CCCGACTGGAAGCGGCGAGTGACGGCAACGCCAATTAATGTGAGTTAGTCACTCATTAG	1200
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
Qy	1261	TCAGGCAGCCATCGGAAGCTGTGTGATGGCTGTGCAGGTGCTAAATCACTGCGATAATTCG	1320
Db	1261	TCAGGCAGCCATCGGAAGCTGTGTGATGGCTGTGCAGGTGCTAAATCACTGCGATAATTCG	1320
Qy	1321	TGTCGCTCAAGGGCGCATCCCGTTCTCGGATAATGTTTTTTTGGCCGACATCAATACGGTT	1380
Db	1321	TGTCGCTCAAGGGCGCATCCCGTTCTCGGATAATGTTTTTTTGGCCGACATCAATACGGTT	1380
Qy	1381	CTGGCAATATCTGAAATGAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
Db	1381	CTGGCAATATCTGAAATGAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
Qy	1441	ATTGTGAGCGGATAACAAATTTTCCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTTCCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500
Qy	1501	GCATTTACCAACAGAGACCATAGATTATGAAATACTGAAGAGTAACTGGTAACTCTGG	1560
Db	1501	GCATTTACCAACAGAGACCATAGATTATGAAATACTGAAGAGTAACTGGTAACTCTGG	1560
Qy	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGGAATTAAGATCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCACAGTT	1680
Db	1621	ACCGGAATTAAGATCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCACAGTT	1680
Qy	1681	CGCGCAACTGGCGATGGCCCTCGACATTAATCTCTGGGCAACAGACCGTTTGGTGGCTAC	1740
Db	1681	CGCGCAACTGGCGATGGCCCTCGACATTAATCTCTGGGCAACAGACCGTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGCAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGCAAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGC CGTACGTTACACCGGAAGCTGATTCGTTTACCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGC CGTACGTTACACCGGAAGCTGATTCGTTTACCCGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTCTCGCAACCCGCAAAAACTCGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTCTCGCAACCCGCAAAAACTCGGAA	1920

RESULT 3

US-09-265-617B-1

US 03 203 017 B 1
; Sequence 1, Application US/09265617B

; Patent No. 6372883

GENERAL INFORMATION:

APPLICANT: Attwood, Michael R.

APPLICANT: Hurst, David N.

APPLICANT: Jones, Philip S.

APPLICANT: Kay, Paul B.

; APPLICANT: Raynham, Tony M.

APPLICANT: Wilson, Francis X.

FILE REFERENCE: 20052 antiviral medications

FILE REFERENCE: 20052 antiviral medicaments
CURRENT APPLICATION NUMBER: US/00/265 5170

; CURRENT APPLICATION NUMBER: US/09/265,617B
 ; CURRENT FILING DATE: 1998-03-10

; CURRENT FILING DATE: 1999-03-10
 ; PRIOR APPLICATION NUMBER: CB9806

;; PRIOR APPLICATION NUMBER: GB9806815.8
; PRIOR FILING DATE: 1998-03-30

; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24

; NUMBER OF SEQ ID NOS: 24 ;
; SOFTWARE: PatentIn Ver. 3.0.5

```

; COM MAKE: FACILITY VEL. 2.0
; SEQ ID NO 1

```

4
2
2
2
4
x
1
0
-

; LENGTH: 7475									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: Vector and									
; OTHER INFORMATION: Gene Fragments									
US-09-265-617B-1									
Query Match 81.8%; Score 2700; DB 3; Length 7475;									
Beat Local Similarity 100.0%; Pred. No. 0;									
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CGGACACCATCAATGGTGC	AAACCTTTTCGGGTATG	GCATGATAGCCCGGAAGA	GA	60			
DB	1	CGGACACCATCAATGGTGC	AAACCTTTTCGGGTATG	GCATGATAGCCCGGAAGA	GA	60			
QY	61	GTCAATTCAGGGTGGTGA	TGTGAACACAGTAACT	TATACGATGTCGACAGTAT	CGCG	120			
DB	61	GTCAATTCAGGGTGGTGA	TGTGAACACAGTAACT	TATACGATGTCGACAGTAT	CGCG	120			
QY	121	GTGTCCTTATCAGACCG	GTTCCTCCCGGTGGTGA	ACCGCCAGCCACGTTT	CTCGGAAAA	180			
DB	121	GTGTCCTTATCAGACCG	GTTCCTCCCGGTGGTGA	ACCGCCAGCCACGTTT	CTCGGAAAA	180			
QY	181	CGCGGAAAAAGTGAAG	CGCGGATGGCGAGCT	GAAATTTACATTTCCCA	ACCGCGTGGCAC	240			
DB	181	CGCGGAAAAAGTGAAG	CGCGGATGGCGAGCT	GAAATTTACATTTCCCA	ACCGCGTGGCAC	240			
QY	241	AACAACTGGCGGCAAA	CAGTCTGTTGCTGATT	GGCGTTGGCCACCTCC	AGTCTGGGCCCTGC	300			
DB	241	AACAACTGGCGGCAAA	CAGTCTGTTGCTGATT	GGCGTTGGCCACCTCC	AGTCTGGGCCCTGC	300			
QY	301	AGCGCCGTTCGCAAA	TGTTCGGCGGATTAAT	CTCGCGCCGATCAACT	TGGGTGGCCAGCG	360			
DB	301	AGCGCCGTTCGCAAA	TGTTCGGCGGATTAAT	CTCGCGCCGATCAACT	TGGGTGGCCAGCG	360			
QY	361	TGGTGGTGTGATGTAG	AAACGAGCGGTGGAAG	CGCTGTAAAGCGCGGT	GTGCAATC	420			
DB	361	TGGTGGTGTGATGTAG	AAACGAGCGGTGGAAG	CGCTGTAAAGCGCGGT	GTGCAATC	420			
QY	421	TTCTCGGCAACGCGT	CAGTGGGCTGATCAT	TAACTATCCGCTGGAT	GACAGGATGCCA	480			
DB	421	TTCTCGGCAACGCGT	CAGTGGGCTGATCAT	TAACTATCCGCTGGAT	GACAGGATGCCA	480			
QY	481	TTGCTGTGGAAGCT	GCCTGCACATAATGT	TTCCGCGTTAATCT	TGATGTCTCTGACC	540			
DB	481	TTGCTGTGGAAGCT	GCCTGCACATAATGT	TTCCGCGTTAATCT	TGATGTCTCTGACC	540			
QY	541	CACCCATCAACAGTAT	TATTTTCTCCCATGA	AGACGGTACGCGACT	TGGCGGTGGAGCAT	600			
DB	541	CACCCATCAACAGTAT	TATTTTCTCCCATGA	AGACGGTACGCGACT	TGGCGGTGGAGCAT	600			
QY	601	TGGTGGCATTTGGGT	CACAGCAAACTCGC	TGTAGCGGCCCATTA	AGTTCTGTCTCGG	660			
DB	601	TGGTGGCATTTGGGT	CACAGCAAACTCGC	TGTAGCGGCCCATTA	AGTTCTGTCTCGG	660			
QY	661	CGCGTCTGGCTGTGG	CTGGCTGAATAATCT	CACTCGCAATCAAA	TTACGCGCATAG	720			
DB	661	CGCGTCTGGCTGTGG	CTGGCTGAATAATCT	CACTCGCAATCAAA	TTACGCGCATAG	720			
QY	721	CGGAACGGGAAGGCA	CTGGATGCGATGTCG	GGTTTCAACAAAC	CAATGCAAAATGCTGA	780			
DB	721	CGGAACGGGAAGGCA	CTGGATGCGATGTCG	GGTTTCAACAAAC	CAATGCAAAATGCTGA	780			
QY	781	ATGAGGGCATCGT	TCCCACTGCGATGCT	GGTTGCCAACGAT	CAGATGGCGCTGGGGCA	840			
DB	781	ATGAGGGCATCGT	TCCCACTGCGATGCT	GGTTGCCAACGAT	CAGATGGCGCTGGGGCA	840			
QY	841	TGCGGCCATTACCGA	GTCCGGCTGCGGTT	GGTGGGATCTCCG	TAGTGGGATAGC	900			
DB	841	TGCGGCCATTACCGA	GTCCGGCTGCGGTT	GGTGGGATCTCCG	TAGTGGGATAGC	900			
QY	901	ACGATACCGAAGAC	AGCTCATGTTATAT	CCCGCGTTAAC	CCCATCAACAGGATTTTC	960			

DB	901	ACGATACCGAAGACAGCT	CACTGTTATATCCCGCG	TTAAACCAACATCAAA	CAGGATTTTC	960			
QY	961	GCCTGCTGGGCAAAAC	CAGCGTGGACCGCTT	GTGCAACTCTCTCAG	GGCCAGGCGTGA	1020			
DB	961	GCCTGCTGGGCAAAAC	CAGCGTGGACCGCTT	GTGCAACTCTCTCAG	GGCCAGGCGTGA	1020			
QY	1021	AGGGCAATCAGCTGT	TGCGCGTCTCACTG	TGTGAAGAAACCA	CCCTGGCGCCCAATA	1080			
DB	1021	AGGGCAATCAGCTGT	TGCGCGTCTCACTG	TGTGAAGAAACCA	CCCTGGCGCCCAATA	1080			
QY	1081	CGCAAAACCGCTCT	CCCCCGCGTGGCGG	ATTCATTAATGCAG	CTGCACAGAGTTT	1140			
DB	1081	CGCAAAACCGCTCT	CCCCCGCGTGGCGG	ATTCATTAATGCAG	CTGCACAGAGTTT	1140			
QY	1141	CCGACTGGAAGGGG	CAGTGAGCGCAAC	CGCAATTAATGTG	AGTTAGTCACTCAT	1200			
DB	1141	CCGACTGGAAGGGG	CAGTGAGCGCAAC	CGCAATTAATGTG	AGTTAGTCACTCAT	1200			
QY	1201	GCACAAATTCATG	TGTGACAGCTTAT	CATGCACTGCA	CGGTGCACCAATG	1260			
DB	1201	GCACAAATTCATG	TGTGACAGCTTAT	CATGCACTGCA	CGGTGCACCAATG	1260			
QY	1261	TCAGCAGCCATCG	GAAGCTGTGATG	CTGTGCTGAG	GTGCTGTAATCA	1320			
DB	1261	TCAGCAGCCATCG	GAAGCTGTGATG	CTGTGCTGAG	GTGCTGTAATCA	1320			
QY	1321	TGTGCTCAAGCG	CACCTCCCGTCT	TGATTAATGTTT	TGCGCGACATCA	1380			
DB	1321	TGTGCTCAAGCG	CACCTCCCGTCT	TGATTAATGTTT	TGCGCGACATCA	1380			
QY	1381	CTGGCAAAATTT	CTGAAATGAGCT	GTGCAATTAAT	CATCGGCTCGTAT	1440			
DB	1381	CTGGCAAAATTT	CTGAAATGAGCT	GTGCAATTAAT	CATCGGCTCGTAT	1440			
QY	1441	ATTGTGAGCGAT	TAACAAATTTCA	CAGAGAACCGC	AGTCCGTTAGGTTT	1500			
DB	1441	ATTGTGAGCGAT	TAACAAATTTCA	CAGAGAACCGC	AGTCCGTTAGGTTT	1500			
QY	1501	GCACCTTCAAC	CAAGGACCATAG	ATTATGAAACT	GAAAGGTAATACT	1560			
DB	1501	GCACCTTCAAC	CAAGGACCATAG	ATTATGAAACT	GAAAGGTAATACT	1560			
QY	1561	ATTAACCGCGAT	TAAGGCTATAA	CGGTCTCGTGA	AGTTCGGAAGAT	1620			
DB	1561	ATTAACCGCGAT	TAAGGCTATAA	CGGTCTCGTGA	AGTTCGGAAGAT	1620			
QY	1621	ACCGGAATTAAG	TAAAGTCAACCG	TTCGAGTAACT	GGAAGAAATTC	1680			
DB	1621	ACCGGAATTAAG	TAAAGTCAACCG	TTCGAGTAACT	GGAAGAAATTC	1680			
QY	1681	GGCGCACTGG	CGATGGCCCTG	ACATTAATCTT	CTGGGCAACAG	1740			
DB	1681	GGCGCACTGG	CGATGGCCCTG	ACATTAATCTT	CTGGGCAACAG	1740			
QY	1741	GCTCAATTCG	CGCTGTGGCTGA	AAATCAACCCG	GCAAAAGCGTTC	1800			
DB	1741	GCTCAATTCG	CGCTGTGGCTGA	AAATCAACCCG	GCAAAAGCGTTC	1800			
QY	1801	CGTTTACCT	GGGATGGCGTAC	GTTAACCGC	AGCTGATTC	1860			
DB	1801	CGTTTACCT	GGGATGGCGTAC	GTTAACCGC	AGCTGATTC	1860			
QY	1861	GAAGGTTAT	CGCTGATTTATA	CAAGATCTG	TGCGCAACCC	1920			
DB	1861	GAAGGTTAT	CGCTGATTTATA	CAAGATCTG	TGCGCAACCC	1920			
QY	1921	GAGATCCCG	CGCTGGATAA	AGAACTGAA	AGCAAGGTAAG	1980			
DB	1921	GAGATCCCG	CGCTGGATAA	AGAACTGAA	AGCAAGGTAAG	1980			
QY	1981	CTGCAAGAAC	CGTACTTCA	CTGGCGCTG	ATTGCTGCTAG	2040			
DB	1981	CTGCAAGAAC	CGTACTTCA	CTGGCGCTG	ATTGCTGCTAG	2040			

Db 1981 CTGCAAGAACCGTACTTCACTCGCGCGTGAATTGCTGTGACGCGGGGTATCGCTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAAAGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAAAGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGACCTGATTAATAAACAACACATGAATGCGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGACCTGATTAATAAACAACACATGAATGCGACACACCGATTAC 2160
QY 2161 TCATCGCAAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCGCTGG 2220
Db 2161 TCATCGCAAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACACAGCAAGTGAATTATGTTGTAACGGTACTGCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACAGCAAGTGAATTATGTTGTAACGGTACTGCGACCTTC 2280
QY 2281 AAGGTCACACATCAACACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACACATCAACACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGGCGTGTGAGCGTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGGCGTGTGAGCGTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGTTATGCGCCACCATGGAACAGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATGCGCCACCATGGAACAGCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACAATCCGCGAGATGTCGCTTCTGTTGATGCGTGGTGTGAGCGTGAAGTCTT 2580
Db 2521 CCGAACAATCCGCGAGATGTCGCTTCTGTTGATGCGTGGTGTGAGCGTGAAGTCTT 2580
QY 2581 GCGAGCGTGTGAGTGTGATGAGCGCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCGAGCGTGTGAGTGTGATGAGCGCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
QY 2641 AACACACAAACAATAACAATAACAACACCTCGGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACACACAAACAATAACAATAACAACACCTCGGGATCGAGGAGGATTCAGAAATTC 2700

RESULT 4

US-09-027-169-3/c
; Sequence 3, Application US/09027169
; Patent No. 6420524
; GENERAL INFORMATION:
; APPLICANT: CRAIG, NANCY L
; TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
; TITLE OF INVENTION: ATP-DEPENDENT TRANSCRIPTION PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Anne Brown (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave.
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/027,169
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Brown, Anne
; REGISTRATION NUMBER: 36,463
; REFERENCE/DOCKET NUMBER: 5789-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2205
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "pEM delta R.adj to 1"
; US-09-027-169-3

Query Match 43.0%; Score 1417.6; DB 3; Length 5926;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 9; Indels 12; Gaps 2;

QY 3 GACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGGAAGAGT 62
Db 1564 GACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGGAAGAGT 1505
QY 63 CAATTCAAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATGCCGGT 122
Db 1504 CAATTCAAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATGCCGGT 1445
QY 123 GTCTTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCAGCTTTCTGCGAAAAACG 182
Db 1444 GTCTTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCAGCTTTCTGCGAAAAACG 1385
QY 183 CGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAAA 242
Db 1384 CGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAAA 1325
QY 243 CAATGCGCGGCAACAGTGTCTGCTGATGGCGGTGGCAACCTCCAGTCTGGGCCCTGCAC 302
Db 1324 CAATGCGCGGCAACAGTGTCTGCTGATGGCGGTGGCAACCTCCAGTCTGGGCCCTGCAC 1265
QY 303 GCGCGTTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCAGCGTG 362
Db 1264 GCGCGTTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCAGCGTG 1205
QY 363 GTGCTGTGATGTGTAGAAAGCGCGCTGGAAGCGCTGTAAAGCGCGGTGGCAAACTCTT 422
Db 1204 GTGCTGTGATGTGTAGAAAGCGCGCTGGAAGCGCTGTAAAGCGCGGTGGCAAACTCTT 1145
QY 423 CTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGGATGCCATT 482
Db 1144 CTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGGATGCCATT 1085
QY 483 GCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTATTTCTTGATGTCTCTGACCAAGACA 542
Db 1084 GCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTATTTCTTGATGTCTCTGACCAAGACA 1025
QY 543 CCCATCAACAGATTAATTTCTCCCATGAAGCGGTACCGGCTGGCGCTGGAGCATCTG 602
Db 1024 CCCATCAACAGATTAATTTCTCCCATGAAGCGGTACCGGCTGGCGCTGGAGCATCTG 965
QY 603 GTGCGAATGGGTACCAAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTCTCTCTCGGG 662
Db 964 GTGCGAATGGGTACCAAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTCTCTCTCGGG 905
QY 663 CGTCTCGGCTCGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCGGATAGCG 722
Db 904 CGTCTCGGCTCGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCGGATAGCG 845
QY 723 GAACGGGAAGGCGACTGGAGTGCATGTCCGGTCTTTCACAAACCAATGCAAAATGCTGAAT 782
Db 844 GAACGGGAAGGCGACTGGAGTGCATGTCCGGTCTTTCACAAACCAATGCAAAATGCTGAAT 785
QY 783 GAGGCGATCGTTCCTCCACTCGGATGCTGTTGCGCAACGATCAGATGGCGCTGGGCGCAATG 842

Db 784 GAGGCGATCGTTCACATGCGATGCTGGTGGCAACGATCAGATGGCGCTGGGCGCAATG 725
Qy 843 CGCGCATTTACCGAGTCCGGGCTGGCGCTGGTGGGATATCTCGGTAGTGGGATACGAC 902
Db 724 CGCGCATTTACCGAGTCCGGGCTGGCGCTGGTGGGATATCTCGGTAGTGGGATACGAC 665
Qy 903 GATACCGAAGACAGCTCATGTTATATCCGGCGTTAAACCATCAACAGAGATTTTCG 962
Db 664 GATACCGAAGACAGCTCATGTTATATCCGGCGTTCAACCATCAACAGAGATTTTCG 605
Qy 963 CTGCTGGGGCAAAACAGAGTGGACCGTCTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 604 CTGCTGGGGCAAAACAGAGTGGACCGTCTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 545
Qy 1023 GGCATCAGCTGTTGCCCGTCTCAGTGGTGAAGAAAGAAACACACCTGGCGCCCAATACG 1082
Db 544 GGCATCAGCTGTTGCCCGTCTCAGTGGTGAAGAAAGAAACACACCTGGCGCCCAATACG 485
Qy 1083 CAAACCGCTCTCCCGCGCTGTCAGTGGCCGATTCATTAATGCGAGCTGGCAGCAGGTTTCC 1142
Db 484 CAAACCGCTCTCCCGCGCTGTCAGTGGCCGATTCATTAATGCGAGCTGGCAGCAGGTTTCC 425
Qy 1143 CGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGCGAGTGGCAGCAGGTTTCC 1202
Db 424 CGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGCGAGTGGCAGCAGGTTTCC 376
Qy 1203 ACAATTCATGTTTGACAGCTTATCATCGACTGCAACGCTGACCAATGCTTCTGGCGTC 1262
Db 375 GAATGATCTGGTTGACAGCTTATCATCGACTGCAACGCTGCAACGCTTCTGGCGTC 316
Qy 1263 AGGAGCCATCGGAAGCTGTTGATGGCTGTCAGGTCGTAATCACTGCATAATTCGTG 1322
Db 315 AGGAGCCATCGGAAGCTGTTGATGGCTGTCAGGTCGTAATCACTGCATAATTCGTG 256
Qy 1323 TCGCTCAAGGGCGACTCCGCTCGGATAGTGTGTTTGGCGCGACATCAACGGTTCT 1382
Db 255 TCGCTCAAGGGCGACTCCGCTCGGATAGTGTGTTTGGCGCGACATCAACGGTTCT 196
Qy 1383 GGCATATTCGAAATGAGCTGTTGACAAATTAATCAT- CGGCTCGTATATGTTGGAA 1441
Db 195 GGCATATTCGAAATGAGCTGTTGACAAATTAATCATCCGGCTCGTATATGTTGGAA 136
Qy 1442 TTGTGAGCGGATAACAATTTACAGAGGAAACAGCC 1477
Db 135 TTGTGAGCGGATAACAATTTACAGAGGAAACAGAC 100

RESULT 5

US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE2
US-09-640-882-2

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
Qy 3 GACACCATCGAATGGTGCMAAACCTTTTCGGGTATGGCATATAGCGCCCGGAGAGAGT 62
Db 866 GACACCATCGAATGGTGCMAAACCTTTTCGGGTATGGCATATAGCGCCCGGAGAGAGT 925
Qy 63 CAATTCAGGGTGGTGAATGTGAACACAGTAACAGTTATACGATGTCGAGAGATGTCGCGT 122
Db 926 CAATTCAGGGTGGTGAATGTGAACACAGTAACAGTTATACGATGTCGAGAGATGTCGCGT 985
Qy 123 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGCGCCAGCTTCTTCGCAAAACG 182
Db 986 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGCGCCAGCTTCTTCGCAAAACG 1045
Qy 183 CGGAAAAAGTGAAGCGGCGATCGCGAGCTGAATTAATTCCTCAACCGGTCGACAA 242
Db 1046 CGGAAAAAGTGAAGCGGCGATCGCGAGCTGAATTAATTCCTCAACCGGTCGACAA 1105
Qy 243 CAATCGCGGCAACAGTCTGCTGATGCGGTGGCACTCCAGTCTGGCCCTGAC 302
Db 1106 CAATCGCGGCAACAGTCTGCTGATGCGGTGGCACTCCAGTCTGGCCCTGAC 1165
Qy 303 CGCGCTCGCAAAATGTCGCGCGATTAATTCCTCGCGCGATCACTGGGTGCCAGCGT 362
Db 1166 CGCGCTCGCAAAATGTCGCGCGATTAATTCCTCGCGCGATCACTGGGTGCCAGCGT 1225
Qy 363 GTGCTGTCGATGTAAGACGCGGTGCAAGCTCTGAAGCGGCGGTGCACAACTCT 422
Db 1226 GTGCTGTCGATGTAAGACGCGGTGCAAGCTCTGAAGCGGCGGTGCACAACTCT 1285
Qy 423 CTCGCGCAACCGCTCAGTGGCTGATCATTAACATATCCGTGGATGACAGAGATGCCATT 482
Db 1286 CTCGCGCAACCGCTCAGTGGCTGATCATTAACATATCCGTGGATGACAGAGATGCCATT 1345
Qy 483 GCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTATTTCTTGATGTCCTGACAGACA 542
Db 1346 GCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTATTTCTTGATGTCCTGACAGACA 1405
Qy 543 CCCATCAACAGTATTTCTCCCATGAAGAGTACGCGACTGGGCGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTTCTCCCATGAAGAGTACGCGACTGGGCGTGGAGCATCTG 1465
Qy 603 GTCGATTGGGTCAACAGCAAAATCGCGTGTAGCGGCCCAATTAAGTTCTGTCTCGCG 662
Db 1466 GTCGATTGGGTCAACAGCAAAATCGCGTGTAGCGGCCCAATTAAGTTCTGTCTCGCG 1525
Qy 663 GCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 1526 GCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAACGGGAAGCGACTGGAGTGCATGTCGGTTCACAAACCCATCAAAATGCTGAAT 782
Db 1586 GAACGGGAAGCGACTGGAGTGCATGTCGGTTCACAAACCCATCAAAATGCTGAAT 1645
Qy 783 GAGGCGATCGTTCACCATGCGATGCTGTTGCGCAACGATCAGATGCGCTGGGCGCAATG 842
Db 1646 GAGGCGATCGTTCACCATGCGATGCTGTTGCGCAACGATCAGATGCGCTGGGCGCAATG 1705
Qy 843 CGCGCATTTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACGAC 902
Db 1706 CGCGCATTTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGAGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGAGATTTTCG 1825
Qy 963 CTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1022
Db 1826 CTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1885
Qy 1023 GGCATCAGCTGTTGCGCGTCTCAGTGGTGAAGAAAGAAACCAACCTGGCGCCCAATACG 1082

1886 GGCATATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAACACCTGCGCCCAATAGC 1945
1083 CAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGACGTGGCAGCAGAGTTTCC 1142
1946 CAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGACGTGGCAGCAGAGTTTCC 2005
1143 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGCTAGCTCACTCAATTAGC 1202
2006 CGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGCTAG--CGGAAATTGATC 2063
1203 ACAATCTCATGTTGACAGCTTATCATGCACTGCAAGTGGCAGCAATGCTTCTGCGCTC 1262
2064 TGAATCTCATGTTGACAGCTTATCATGCACTGCAAGTGGCAGCAATGCTTCTGCGCTC 2123
1263 AGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTGGTAAATCACTGCATTAATTCGTG 1322
2124 AGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTGGTAAATCACTGCATTAATTCGTG 2183
1323 TCGCTCAAGGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAAGGTTCT 1382
2184 TCGCTCAAGGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAAGGTTCT 2243
1383 GGCATAATTTCT 1394
2244 GGCATAATTTCT 2255

RESULT 6
US-09-640-882-3
; Sequence 3, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE3
US-09-640-882-3

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
3 GACACCATCGAATCGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGAGT 62
866 GACACCATCGAATCGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGAGT 925
63 CAATTACAGGTGGTGAATGTGAACCAAGTAACGTTATACGATGCGAGAGTATGCCGTT 122
926 CAATTACAGGTGGTGAATGTGAACCAAGTAACGTTATACGATGCGAGAGTATGCCGTT 985
123 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACGTTTCTCGGAAACG 182
986 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACGTTTCTCGGAAACG 1045
183 CGGAAAAAGTGGNAGCGGCGATGGCGAGTGAATTTACATTCCTCCACCGCGTGGCACA 242
1046 CGGAAAAAGTGGNAGCGGCGATGGCGAGTGAATTTACATTCCTCCACCGCGTGGCACA 1105

243 CAACCTGGCGGCAACAGTGGTCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGCGAC 302
1106 CAACCTGGCGGCAACAGTGGTCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGCGAC 1165
303 GCGCGCTGCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAATCTGGGTGGCAGCGTG 362
1166 GCGCGCTGCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAATCTGGGTGGCAGCGTG 1225
363 GTGTGTGCGATGTTAGTAAACGAGCGGCTGAAAGCCTGTAAAGCGGCGGTGCAAAATCTTT 422
1226 GTGTGTGCGATGTTAGTAAACGAGCGGCTGAAAGCCTGTAAAGCGGCGGTGCAAAATCTTT 1285
423 CTGCGCAACGCGTCACTGGGCTGATCAATTAATCTCGCGTGGATGACGAGATGCCATT 482
1286 CTGCGCAACGCGTCACTGGGCTGATCAATTAATCTCGCGTGGATGACGAGATGCCATT 1345
483 GCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTCTGACCAACA 542
1346 GCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTCTGACCAACA 1405
543 CCNATCAACAGTATTTTCTCCATGAAGAGCGGTACGCACTGGGCGTGGAGCATCTG 602
1406 CCCATCAACAGTATTTTCTCCATGAAGAGCGGTACGCACTGGGCGTGGAGCATCTG 1465
603 GTGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCATTAAGTCTGTCGCGG 662
1466 GTGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCATTAAGTCTGTCGCGG 1525
663 CGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAGCG 722
1526 CGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAGCG 1585
723 GAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATGCTGAAT 782
1586 GAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATGCTGAAT 1645
783 GAGGCGATCTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAATG 842
1646 GAGGCGATCTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAATG 1705
843 GCGGCCATTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACGAC 902
1706 GCGGCCATTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACGAC 1765
903 GATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGAGATTTTCG 962
1766 GATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGAGATTTTCG 1825
963 CTGCTGGGCAACACGAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGAAG 1022
1826 CTGCTGGGCAACACGAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGAAG 1885
1023 GGCATACAGCTGTTGCCGCTCTCACTGGTGAAGAAACCAACCTCGGCGCCCAATAGC 1082
1886 GGCATACAGCTGTTGCCGCTCTCACTGGTGAAGAAACCAACCTCGGCGCCCAATAGC 1945
1083 CAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGCAAGTGGCAGCAGAGTTTCC 1142
1946 CAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGCAAGTGGCAGCAGAGTTTCC 2005
1143 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGCTAGCTCACTCAATTAGC 1202
2006 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGCTAG--CGCGAAATTGATC 2063
1203 ACAATCTCATGTTGACAGCTTATCATGCACTGCAAGTGGCAGCAATGCTTCTGCGCTC 1262
2064 TGAATCTCATGTTGACAGCTTATCATGCACTGCAAGTGGCAGCAATGCTTCTGCGCTC 2123
1263 AGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTGGTAAATCACTGCATAAATTCGTG 1322
2124 AGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTGGTAAATCACTGCATAAATTCGTG 2183
1323 TCGCTCAAGGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAAGGTTCT 1382

Db	2184	TCGCTCAAGCGCACTCCCGTTGGGATAATGTTTTTGGCCGACATCATAACGGTTCT	2243
Qy	1383	GGCAATATTCT 1394	
Db	2244	GGCAATATTCT 2255	
<p>RESULT 7</p> <p>US-08-778-717-5</p> <p>; Sequence 5, Application US/08778717</p> <p>; Patent No. 6602689</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: UENO, EIICHI</p> <p>; APPLICANT: NOBUYUKI, FUJII</p> <p>; APPLICANT: OKADA, MASAHISA</p> <p>; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN</p> <p>; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR</p> <p>; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE</p> <p>; NUMBER OF SEQUENCES: 21</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,</p> <p>; ADDRESSEE: P.C.</p> <p>; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400</p> <p>; CITY: ARLINGTON</p> <p>; STATE: VA</p> <p>; COUNTRY: USA</p> <p>; ZIP: 22202</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; COMPUTER: IBM PC compatible</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: PatentIn Release #1.0, Version #1.30</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/08/778,717</p> <p>; FILING DATE: 12-DEC-1996</p> <p>; CLASSIFICATION: 536</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: JP 352225/1995</p> <p>; FILING DATE: 28-DEC-1995</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: OBLON, NORMAN F.</p> <p>; REGISTRATION NUMBER: 24,618</p> <p>; REFERENCE/DOCKET NUMBER: 2084-031-0</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: 703-413-3000</p> <p>; TELEFAX: 703-413-2220</p> <p>; INFORMATION FOR SEQ ID NO: 5:</p> <p>; SEQUENCE CHARACTERISTICS:</p> <p>; LENGTH: 4557 base pairs</p> <p>; TYPE: nucleic acid</p> <p>; STRANDEDNESS: double</p> <p>; TOPOLOGY: circular</p> <p>; MOLECULE TYPE: other nucleic acid</p> <p>; ORIGINAL SOURCE:</p> <p>; ORGANISM: E. COLI</p> <p>; STRAIN: BL21 (DE3)</p> <p>; PUBLICATION INFORMATION:</p> <p>; AUTHORS: NOBUYUKI FUJII ET AL,</p> <p>; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED</p> <p>; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF</p> <p>; TITLE: EXPRESSING SAID FUSED PROTEIN</p> <p>; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557</p> <p>US-08-778-717-5</p> <p>Query Match 36.5%; Score 1204; DB 4; Length 4557;</p> <p>Best Local Similarity 100.0%; Pred.No. 0;</p> <p>Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>			
Qy	1	CCGACACATCGAATGGTGCMAAACCTTTTCGCGGTATGGCATGATAGGCCCGGAAGAGA	60
Db	2563	CCGACACATCGAATGGTGCMAAACCTTTTCGCGGTATGGCATGATAGGCCCGGAAGAGA	2622

Db 3703 CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTACTACTATTAG 3762
Qy 1201 GCAC 1204
Db 3763 GCAC 3766

RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerra, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3832 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-148-675A-2

Query Match 36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGGCAAAACCTTCCGGTATGGCATGATAGCGCCCGGAAGA 60
Db 3 CCGACACCATCGAATGGCGCAAAACCTTCCGGTATGGCATGATAGCGCCCGGAAGA 62

Qy 61 GTCAATTTCAGGGTGGTGAATGTGAAACAGTAACGTTTATACGATGTCGACAGATATGCCG 120
Db 63 GTCAATTTCAGGGTGGTGAATGTGAAACAGTAACGTTTATACGATGTCGACAGATATGCCG 122

Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAAACAGTAACGTTTATACGATGTCGACAGATATGCCG 120
Db 123 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAAACAGTAACGTTTATACGATGTCGACAGATATGCCG 122

Qy 181 CGCGGGAAGTGAAGCGCGCATGGCGAGTGAATTCATTCCTCCAAACCGCGTGGCAC 240
Db 183 CGCGGGAAGTGAAGCGCGCATGGCGAGTGAATTCATTCCTCCAAACCGCGTGGCAC 242

Qy 241 AACCACTGGCGGGCAACAGTCGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300

Db 243 AACCACTGGCGGGCAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 302
Qy 301 ACGCCGCGTCCCAAAATTTGTCGGCGGATTTAAATCTCGCGCGGATCAATCTGGTCCGACGC 360
Db 303 ACGCCGCGTCCCAAAATTTGTCGGCGGATTTAAATCTCGCGCGGATCAATCTGGTCCGACGC 362
Qy 361 TGGTGGTGTGATCGTAGAAGCGGCGTCCGAAGCCTCTGAAGCGGCTGTAAGCGCGGTGCACATC 420
Db 363 TGGTGGTGTGATCGTAGAAGCGGCGTCCGAAGCCTCTGAAGCGGCGGTGCACATC 422
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATCGACAGGATGCCA 480
Db 423 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATCGACAGGATGCCA 482
Qy 481 TTGCTGTGGAGCTGCGCTGCACTAAATGTTCCGGGCTTATTTCTTGATGTCCTCGACAGA 540
Db 483 TTGCTGTGGAGCTGCGCTGCACTAAATGTTCCGGGCTTATTTCTTGATGTCCTCGACAGA 542
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 543 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 602
Qy 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGCTCGG 660
Db 603 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGCTCGG 662
Qy 661 CGGCTCTGCGCTGCGCTGGCTGGCGGATTAATCTCTACTCGCAATCAAAATTCAGCCGATG 720
Db 663 CGGCTCTGCGCTGCGCTGGCTGGCGGATTAATCTCTACTCGCAATCAAAATTCAGCCGATG 722
Qy 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTCCTCAACAAACCATCAAAATGCTGA 780
Db 723 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTCCTCAACAAACCATCAAAATGCTGA 782
Qy 781 ATGAGGCGATCGTTCCCATCGCATGTCGATGTCGCAACGATCAGATGGCGCTGGCGCAA 840
Db 783 ATGAGGCGATCGTTCCCATCGCATGTCGATGTCGCAACGATCAGATGGCGCTGGCGCAA 842
Qy 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGATGTCGGGATATCTCGGTAGTGGGATACG 900
Db 843 TGGCGGCCATTACCGAGTCCGGGCTGGCGGATGTCGGGATATCTCGGTAGTGGGATACG 902
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGATTTTC 960
Db 903 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGATTTTC 962
Qy 961 GCCTGCTGGGGCAACACGAGTGGACCGCTTGTGTCGAACTCTCTCAGGGCCAGCGGTGA 1020
Db 963 GCCTGCTGGGGCAACACGAGTGGACCGCTTGTGTCGAACTCTCTCAGGGCCAGCGGTGA 1022
Qy 1021 AGGGCAATCAGCTGTTTGGCCGCTCTCACTGTTGAAAAGAAAACCCCTGGCGCCCAATA 1080
Db 1023 AGGGCAATCAGCTGTTTGGCCGCTCTCACTGTTGAAAAGAAAACCCCTGGCGCCCAATA 1082
Qy 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATGCGAGTGGCACGACAGGTTT 1140
Db 1083 CGCAAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATGCGAGTGGCACGACAGGTTT 1142
Qy 1141 CCGGACTGGAAGCGGGCGAGTGGCGCAACGCAATTAATGCGAGTGGCTCACTCATTTAG 1200
Db 1143 CCGGACTGGAAGCGGGCGAGTGGCGCAACGCAATTAATGCGAGTGGCTCACTCATTTAG 1202
Qy 1201 GCAC 1204
Db 1203 GCAC 1206

RESULT 9
US-08-487-283A-18/c
; Sequence 18, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.

APPLICANT: Matis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Squinto, Stephen P.
APPLICANT: Thomas, Thomas C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
TITLE OF INVENTION: OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: Seth A. Fidel
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Cetrizis 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)776-1790
TELEFAX: (203)772-3655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: pET Trc SOS/NI
DESCRIPTION: prokaryotic expression vector
US-08-487-283A-18

Query Match 36.3%; Score 1199.2; DB 3; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGGACACCATCGAATGGTGCAGAACTTTTCGGGGTATGGCATGATAGCGCCGGAGAGA 60
DB 4810 CGGACACCATCGAATGGCGCAAACTTTTCGGGGTATGGCATGATAGCGCCGGAGAGA 4751
QY 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAACGTATTACCATGTGCGAGAGTATGCCG 120
DB 4750 GTCAATTCAGGGTGTGAATGTGAACACAGTAACGTATTACCATGTGCGAGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGTTTCCCGGGTGTGAACACAGCGACGCCACGTTTCTCGAAAA 180
DB 4690 GTGTCTCTTATCAGACCGTTTCCCGGGTGTGAACACAGCGACGCCACGTTTCTCGAAAA 4631
QY 181 CGCGGAAAAAGTGAACGGCGATGGCGAGCTGAATTCACATCCCAACCGCGTGGCAC 240
DB 4630 CGCGGAAAAAGTGAACGGCGATGGCGAGCTGAATTCACATCCCAACCGCGTGGCAC 4571
QY 241 AACAACTGCGCGGCAAAACAGTCTGCTGATTTGGGGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 4570 AACAACTGCGCGGCAAAACAGTCTGCTGATTTGGGGTTGCCACCTCCAGTCTGGCCCTGC 4511
QY 301 AGCGCGCGTGCAGAAATTTGTCGGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 360

RESULT 10

PCI-US96-05611A-21/c
; Sequence 21, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Lenardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott

Db 4510 ACGGCCGTGCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCGCAGC 4451
QY 361 TGGTGGTGTGATGTAGAACGAAAGCGCGTCTGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 4450 TGGTGGTGTGATGTAGAACGAAAGCGCGTCTGAAGCCTGTAAAGCGCGGTGCAACAATC 4391
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACGAGATGCCA 480
Db 4390 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACGAGATGCCA 4331
QY 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db 4330 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 4271
QY 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 4270 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 4211
QY 601 TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGGCCATTAAAGTTCTGTCTCG 660
Db 4210 TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGGCCATTAAAGTTCTGTCTCG 4151
QY 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4150 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 4090 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCGCAACGATGCGGCTGGCGGCAAA 840
Db 4030 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCGCAACGATGCGGCTGGCGGCAAA 3971
QY 841 TGGCGGCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAGC 900
Db 3970 TGGCGGCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAGC 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAAAACAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAAAACAGGATTTTC 3851
QY 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3850 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 3791
QY 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Db 3790 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 3731
QY 1081 CGCAAAACCGCTCTCCCGCGCGTGGCGGATTCAATTAATGAGCTGGCAGCAGAGTTT 1140
Db 3730 CGCAAAACCGCTCTCCCGCGCGTGGCGGATTCAATTAATGAGCTGGCAGCAGAGTTT 3671
QY 1141 CCCGACTGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 3670 CCCGACTGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607

; APPLICANT: Nye, Steven H.
 ; APPLICANT: Pelfrey, Clara M.
 ; APPLICANT: Squinto, Stephen P.
 ; APPLICANT: Wilkins, James A.
 ; TITLE OF INVENTION: Modified Myelin Protein Molecules
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Maurice M. Klee
 ; STREET: 1951 Burr Street
 ; CITY: Fairfield
 ; STATE: Connecticut
 ; COUNTRY: USA
 ; ZIP: 06430

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
 ; COMPUTER: Macintosh Centris 610
 ; OPERATING SYSTEM: Sysvtem 7
 ; SOFTWARE: Microsoft Word 6.0.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/05611A
 ; FILING DATE: 02-MAY-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/431,644
 ; FILING DATE: May 2, 1995
 ; APPLICATION NUMBER: 08/431,648
 ; FILING DATE: May 2, 1995
 ; APPLICATION NUMBER: 08/482,114
 ; FILING DATE: June 7, 1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Klee, Maurice M.
 ; REGISTRATION NUMBER: 30,399
 ; REFERENCE/DOCKET NUMBER: ALX-129
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (203) 255 1400
 ; TELEFAX: (203) 254 1101
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5248 base pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Double
 ; TOPOLOGY: Circular
 ; MOLECULE TYPE: Other nucleic acid
 ; DESCRIPTION: PET Trc 805/NI
 ; DESCRIPTION: prokaryotic expression vector
 ; PCT-US96-05611A-21

Query Match 36.3%; Score 1199.2; DB 5; Length 5248;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGGTGAACAACTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
 DB 4810 CGGACACCATCGAATGGCGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 4751
 QY 61 GTCAATTACAGGTGTAAGTGAACACAGTAACTGATGCGTATGCGATGCGAGATGATCCG 120
 DB 4750 GTCAATTACAGGTGTAAGTGAACACAGTAACTGATGCGTATGCGATGCGAGATGATCCG 4691
 QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCGAGCGACGCTTTCGCGAAAA 180
 DB 4690 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCGAGCGACGCTTTCGCGAAAA 4631
 QY 181 CGCGGGAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGGTGGCAC 240
 DB 4630 CGCGGGAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGGTGGCAC 4571
 QY 241 AACAACTGGCGGGCAACAGTCTGTTGCTGATGCGTTCACCTCCAGTCTGGCCCTGC 300
 DB 4570 AACAACTGGCGGGCAACAGTCTGTTGCTGATGCGTTCACCTCCAGTCTGGCCCTGC 4511
 QY 301 ACGCGCGTGCAGAAATTGTCCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
 ;

DB 4510 ACGCGCGTGCAGAAATTGTCCGCGGATTAATCTCGCGCGATCAACTCGGTGCCAGCG 4451
 QY 361 TGGTGGTGTGATGATAGAGAGCGCGTTCGAAGCGCTCTAAAGCGCGCGTGCACAAATC 420
 DB 4450 TGGTGGTGTGATGATAGAGAGCGCGTTCGAAGCGCTCTAAAGCGCGCGTGCACAAATC 4391
 QY 421 TTCTCGCGCAACCGGTGATGCGGTGATCACTAATCACTATCGCTGGATGACAGATGCCA 480
 DB 4390 TTCTCGCGCAACCGGTGATGCGGTGATCACTAATCACTATCGCTGGATGACAGATGCCA 4331
 QY 481 TTGCTGTGGAAGCTGCTGCACAAATGTTCCGGGTTAATTTCTTGTGATGTTCTCTCACCAGA 540
 DB 4330 TTGCTGTGGAAGCTGCTGCACAAATGTTCCGGGTTAATTTCTTGTGATGTTCTCTCACCAGA 4271
 QY 541 CACCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGTGGAGCATC 600
 DB 4270 CACCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGTGGAGCATC 4211
 QY 601 TGGTGGCATTTGGGTACACGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
 DB 4210 TGGTGGCATTTGGGTACACGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 4151
 QY 661 CGGTCTGCGTCTGGGTGGGTGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 DB 4150 CGGTCTGCGTCTGGGTGGGTGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
 QY 721 CGGAACGGGAAGCGGTGAGTGCATGTCGGTTCGCGTTTCAACAAACCATCAAAATGCTGA 780
 DB 4090 CGGAACGGGAAGCGGTGAGTGCATGTCGGTTCGCGTTTCAACAAACCATCAAAATGCTGA 4031
 QY 781 ATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 840
 DB 4030 ATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 3971
 QY 841 TGGCGCGCATTCAGGATCGCGGTGCGGTGCGGATATCTCGGTAGTGGGATACG 900
 DB 3970 TGGCGCGCATTCAGGATCGCGGTGCGGTGCGGATATCTCGGTAGTGGGATACG 3911
 QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGATTTTC 960
 DB 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGATTTTC 3851
 QY 961 GCCTGTGGGGCAACACGCGTGGACCGGTGTCGCAACTCTCTCAGGGCCAGCGGTGA 1020
 DB 3850 GCCTGTGGGGCAACACGCGTGGACCGGTGTCGCAACTCTCTCAGGGCCAGCGGTGA 3791
 QY 1021 AGGCAATCAGCTGTTTCCCGGTCTCACTGGTGAAGAAAAACCCACCTGGCGCCCAATA 1080
 DB 3790 AGGCAATCAGCTGTTTCCCGGTCTCACTGGTGAAGAAAAACCCACCTGGCGCCCAATA 3731
 QY 1081 CGCAACCCGCTCTCCCGCGGTGGCGGATTCATTAATGAGCTGGCACACAGGTTT 1140
 DB 3730 CGCAACCCGCTCTCCCGCGGTGGCGGATTCATTAATGAGCTGGCACACAGGTTT 3671
 QY 1141 CCGACTGAAGCGGGCGGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTATTAG 1200
 DB 3670 CCGACTGAAGCGGGCGGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTATTAG 3611
 QY 1201 GCAC 1204
 DB 3610 GCAC 3607

RESULT 11
 US-10-263-103-35/c
 ; Sequence 35, Application US/10263103
 ; Patent No. 6821723
 ; GENERAL INFORMATION:
 ; APPLICANT: AVENTIS PASTEUR
 ; APPLICANT: Chevallier, Michel
 ; APPLICANT: El Habib, Raphaelle
 ; APPLICANT: Krell, Tino
 ; APPLICANT: Sodoyer, Regis

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pET-cer
US-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGTCAAAACCTTTCCGGGTATGCGATAGAGCCCGGAAGAGA 60
DB 4622 CGGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGCGATAGAGCCCGGAAGAGA 4563

QY 61 GTCGAATTCAGGGTGTGAATGTGAACCAAGTAAACGTTATACGATGTGCGAGAGTATGCCG 120
DB 4562 GTCGAATTCAGGGTGTGAATGTGAACCAAGTAAACGTTATACGATGTGCGAGAGTATGCCG 4503

QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCGGACGACCGTTTCTGCGAAAA 180
DB 4502 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCGGACGACCGTTTCTGCGAAAA 4443

QY 181 CGCGGAAAAAGTGAACGGCGGATGGCGGAGCTGAATTACATTCCTCAACCGCTGGGCAC 240
DB 4442 CGCGGAAAAAGTGAACGGCGGATGGCGGAGCTGAATTACATTCCTCAACCGCTGGGCAC 4383

QY 241 AACAACTCGCGGCAAAACAGTGTGCTGATTGGGGTGGCCACCTCCAGTCTGGCCCTGC 300
DB 4382 AACAACTCGCGGCAAAACAGTGTGCTGATTGGGGTGGCCACCTCCAGTCTGGCCCTGC 4323

QY 301 ACGCCCGCTCGCAAAATTTGCGCGCGATTAATCTCGCGCGGATCAATCGGTTGCCAGCG 360
DB 4322 ACGCCCGCTCGCAAAATTTGCGCGCGATTAATCTCGCGCGGATCAATCGGTTGCCAGCG 4263

QY 361 TGGTGGTTCGATGTAGAACGAACGGCGCTCGAAGCCCTGTAAAGCGGCGGTGCAATC 420
DB 4262 TGGTGGTTCGATGTAGAACGAACGGCGCTCGAAGCCCTGTAAAGCGGCGGTGCAATC 4203

QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACACGAGTGC 480
DB 4202 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACACGAGTGC 4143

QY 481 TTGCTGTGGAAGCTGCCTGCACATAATGTTTCGGCGGTTATTTCTGATGTCTTGACCCAGA 540
DB 4142 TTGCTGTGGAAGCTGCCTGCACATAATGTTTCGGCGGTTATTTCTGATGTCTTGACCCAGA 4083

QY 541 CACCCATCAACGATATTATTTCTCCATGAAGACGGTACGGCATCGGCGTGGAGCATC 600
DB 4082 CACCCATCAACGATATTATTTCTCCATGAAGACGGTACGGCATCGGCGTGGAGCATC 4023

QY 601 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
DB 4022 TGGTGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 3963

QY 661 CCGCTCTGCGTCTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 3962 CCGCTCTGCGTCTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 3903

QY 721 CGGAACGGGAACGGCATGGAATGCGATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 780
DB 3902 CGGAACGGGAACGGCATGGAATGCGATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 3843

QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGTGGCGGCA 840
DB 3842 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGTGGCGGCA 3783

QY 841 TGCSCGCCATTACCGAGTCCGGGCTCGCGTGTGTCGGATATCTCGTATGGGATACG 900
DB 3782 TGCSCGCCATTACCGAGTCCGGGCTCGCGTGTGTCGGATATCTCGTATGGGATACG 3723

QY 901 ACGATACCGGAAGACAGCTCATGTTTATATCCCGCGGTTAAACCAACCATCAAAACGAGTTTC 960
DB 3722 ACGATACCGGAAGACAGCTCATGTTTATATCCCGCGGTTAAACCAACCATCAAAACGAGTTTC 3663

QY 961 GCCTGCTGGGCAAAACCAAGCGTGGACCGCTTGTGTCGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 3662 GCCTGCTGGGCAAAACCAAGCGTGGACCGCTTGTGTCGCAACTCTCTCAGGGCCAGCGGTGA 3603

QY 1021 AGGCAATCAGCTGTGTCGGCTCTCACTGTGTAAGAAAGAAAACCAACCTTGGCGCCCAATA 1080
DB 3602 AGGCAATCAGCTGTGTCGGCTCTCACTGTGTAAGAAAGAAAACCAACCTTGGCGCCCAATA 3543

QY 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGGCAGCAGAGGTTT 1140
DB 3542 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGGCAGCAGAGGTTT 3483

QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 3482 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 3423

QY 1201 GCAC 1204
DB 3422 GCAC 3419

RESULT 12
US-08-929-967-1/c
; Sequence 1, Application US/089299967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Rudpert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;


```
Db 4257 TTCTCGCGCAACGGCTCAGTGGCTGATCATTAACATATCCGCTGGATGACCAAGATGCCA 4198
Qy 481 TTGCTGTGGAGCTGCCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 540
Db 4197 TTGCTGTGGAGCTGCCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 4138
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAGAGCTGACCGACTGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTCTCCCATGAGAGCTGACCGACTGGCGGTGGAGCATC 4078
Qy 601 TGGTCCGATTTGGTCCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Db 4077 TGGTCCGATTTGGTCCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 4018
Qy 661 CGCGTCTCGCTGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTCGCTGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCMAAACCACATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCMAAACCACATGCAATGCTGA 3898
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAACAGATGCGCTGGGCGCAA 840
Db 3897 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAACAGATGCGCTGGGCGCAA 3838
Qy 841 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 3778
Qy 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAACAGGATTTTC 960
Db 3777 ACGATACGGAAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAACAGGATTTTC 3718
Qy 961 GCCTGCTGGGCAACACAGCTGAGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGCAACACAGCTGAGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 3658
Qy 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3657 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 3598
Qy 1081 CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
Db 3597 CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 3538
Qy 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGTCAGTTAGTCTCACTCATTAG 1200
Db 3537 CCCGACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGTCAGTTAGTCTCACTCATTAG 3478
Qy 1201 GCAC 1204
Db 3477 GCAC 3474
```

RESULT 14

```
US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
```

```
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGFTGCAAAACCTTTTCGGGTATGAGCATGATGAGCCCGGAAGAGA 60
Db 4677 CCGACACCATCGAATGFTGCAAAACCTTTTCGGGTATGAGCATGATGAGCCCGGAAGAGA 4618
Qy 61 GTCAATTCAGGGTGGTGAATGTAAGAACAGTAAAGTTATACGATGTCGACAGTATGCG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTAAGAACAGTAAAGTTATACGATGTCGACAGTATGCG 4558
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACAGGCCAGCCACGCTTCTCGGAAAA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACAGGCCAGCCACGCTTCTCGGAAAA 4498
Qy 181 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 4438
Qy 241 AACAACTGGCGGGCAACAGTCTGTGATGTTGGGTGCGCACCTCCAGTCTGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAACAGTCTGTGATGTTGGGTGCGCACCTCCAGTCTGGCCCTGC 4378
Qy 301 ACGCGCGCTCCGCAAAATTTCTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCGCAGC 360
Db 4377 ACGCGCGCTCCGCAAAATTTCTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCGCAGC 4318
Qy 361 TGGTGGTTCGATGGTGAAGACGAGCGCGCTGCAAGCTGTAAAGCGCGGTGCAACAATC 420
Db 4317 TGGTGGTTCGATGGTGAAGACGAGCGCGCTGCAAGCTGTAAAGCGCGGTGCAACAATC 4258
Qy 421 TTCTCGCGCAACGGTCACTGAGTGGGTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 480
Db 4257 TTCTCGCGCAACGGTCACTGAGTGGGTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA 4198
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 540
Db 4197 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 4138
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAGAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTTCTCCCATGAGAGCGGTACGCGACTGGGCGTGGAGCATC 4078
Qy 601 TGGTCCGATTTGGTCCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Db 4077 TGGTCCGATTTGGTCCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 4018
Qy 661 CGCGTCTCGCTGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTCGCTGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCMAAACCACATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCMAAACCACATGCAATGCTGA 3898
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAACAGATGCGCTGGGCGCAA 840
Db 3897 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAACAGATGCGCTGGGCGCAA 3838
Qy 841 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 3778
```


QY 901 ACATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAACAGGATTTTC 960
Db 3777 ACATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA 3658
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCACTGTTGAAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTGCCCGCTCACTGTTGAAAGAAAGAAACCAACCTGGCGCCCAATA 3598
QY 1081 CGCAAAACCGCTCTCCCCCGGCTTGGCGATTCAATTAATGAGCTGGCAGCAGGTTT 1140
Db 3597 CGCAAAACCGCTCTCCCCCGGCTTGGCGATTCAATTAATGAGCTGGCAGCAGGTTT 3538
QY 1141 CCGGACTGAAAGCGGGCAGTGAAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG 1200
Db 3537 CCGGACTGAAAGCGGGCAGTGAAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 15
US-09-614-124B-785/c
; Sequence 785, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGAACACCTTTCGGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 4677 CCGACACCATCGAATGGCGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGAAGAGA 4618
QY 61 GTCAATTGAGGCTGGTGAATGTGAACACAGTAAACAGTTATACGATGTCGAGAGTATGCCG 120
Db 4617 GTCAATTGAGGCTGGTGAATGTGAACACAGTAAACAGTTATACGATGTCGAGAGTATGCCG 4558
QY 121 GTGTCCTTATCAGACCGCTTTCGGCGTGTGAACACAGGCGAGCCAGCTTCTGCGAAAA 180
Db 4557 GTGTCCTTATCAGACCGCTTTCGGCGTGTGAACACAGGCGAGCCAGCTTCTGCGAAAA 4498
QY 181 CCGGGGAAAAAGTGAAGCGCGAGTGGCGAGTGAATTACATTCCTCCAAACCGCGTGGCAC 240
Db 4497 CCGGGGAAAAAGTGAAGCGCGAGTGGCGAGTGAATTACATTCCTCCAAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCCACTCCAGCTCTGGCCCTGC 300

Db 4437 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCCACTCCAGCTCTGGCCCTGC 4378
QY 301 ACGGCGCGTCCCAAAATTTGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGTGCCGACG 360
Db 4377 ACGGCGCGTCCCAAAATTTGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGTGCCGACG 4318
QY 361 TGGTGGTGTGATCGGTGAAGAGCGCGCTGCAAGCGCTGTAAGCGCGCGGTGCACAATC 420
Db 4317 TGGTGGTGTGATCGGTGAAGAGCGCGCTGCAAGCGCTGTAAGCGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
Db 4257 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 4198
QY 481 TTGCTGTGGAAGCTGCTGCACCTAATGTTCCGGCGTTAATTTCTTGATGTCTCTGACCAGA 540
Db 4197 TTGCTGTGGAAGCTGCTGCACCTAATGTTCCGGCGTTAATTTCTTGATGTCTCTGACCAGA 4138
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 4078
QY 601 TGGTGGCATTTGGGTCACCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGCTCGG 660
Db 4077 TGGTGGCATTTGGGTCACCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGCTCGG 4018
QY 661 CGGCTCTGCGCTGCGGTGGCTGATCAATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGGCTCTGCGCTGCGGTGGCTGATCAATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGCGGCTGAGTGGCATGTCGGTTCGCGTTTCAACAAACCATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGCGGCTGAGTGGCATGTCGGTTCGCGTTTCAACAAACCATGCAATGCTGA 3898
QY 781 ATGAGGCGCATGTTCCCACTGCGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Db 3897 ATGAGGCGCATGTTCCCACTGCGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 3838
QY 841 TGGCGGCATTAACGAGTCCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 900
Db 3837 TGGCGGCATTAACGAGTCCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 3778
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTACCACCATCAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTACCACCATCAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACCGGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGCGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACCGGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGCGTGA 3658
QY 1021 AGGGCAATCAGCTGTTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 3598
QY 1081 CGCAAAACCGCTCTCCCCCGGCTTGGCGGATTTCAATTAATGAGCTGGCAGCAGGTTT 1140
Db 3597 CGCAAAACCGCTCTCCCCCGGCTTGGCGGATTTCAATTAATGAGCTGGCAGCAGGTTT 3538
QY 1141 CCGGACTGAAAGCGGGCAGTGAAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG 1200
Db 3537 CCGGACTGAAAGCGGGCAGTGAAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

Search completed: August 3, 2005, 02:32:35
Job time : 356.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds

(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-18

Perfect score: 3300

Sequence: 1 ccgacacatcgatgtgc.....acgacgttcggaactagct 3300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PTCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PTCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3300	100.0	3300	10	US-09-765-555-18	Sequence 18, Appl
2	3232.8	98.0	3300	10	US-09-765-555-17	Sequence 17, Appl
3	3226.4	97.8	3300	10	US-09-765-555-14	Sequence 15, Appl
4	3226.4	97.8	3300	10	US-09-765-555-15	Sequence 15, Appl
5	3216.8	97.5	3300	10	US-09-765-555-16	Sequence 16, Appl
6	2705.8	82.0	7373	18	US-10-257-384A-1	Sequence 1, Appl
7	2705.8	82.0	8101	17	US-10-149-472-5	Sequence 5, Appl

8	2701	81.8	5648	19	US-10-343-859-8	Sequence 8, Appl
9	2698	81.8	9191	19	US-10-343-859-9	Sequence 9, Appl
10	2690.4	81.5	7112	18	US-10-263-153-40	Sequence 40, Appl
11	2690.4	81.5	7259	18	US-10-263-153-30	Sequence 30, Appl
12	2690.4	81.5	7322	18	US-10-263-153-35	Sequence 35, Appl
13	2690.4	81.5	7352	18	US-10-263-153-25	Sequence 25, Appl
14	2690.4	81.5	7370	18	US-10-263-153-20	Sequence 20, Appl
15	2690.4	81.5	7370	18	US-10-263-153-61	Sequence 61, Appl
16	2690.4	81.5	7370	18	US-10-263-153-66	Sequence 66, Appl
17	2690.4	81.5	7370	18	US-10-263-153-71	Sequence 71, Appl
18	2690.4	81.5	7403	18	US-10-263-153-15	Sequence 15, Appl
19	2690.4	81.5	7442	18	US-10-263-153-10	Sequence 10, Appl
20	2690.4	81.5	7478	18	US-10-263-153-3	Sequence 3, Appl
21	2636.8	79.9	6806	10	US-09-908-943A-194	Sequence 194, App
22	2636.8	79.9	6806	20	US-10-801-487-194	Sequence 194, App
23	2636.8	79.9	6806	20	US-10-801-938-194	Sequence 194, App
24	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
25	2636.8	79.9	6806	20	US-10-801-486-194	Sequence 194, App
26	2636.8	79.9	6806	21	US-10-801-493-194	Sequence 194, App
27	2605.4	79.0	7553	18	US-10-263-153-7	Sequence 7, Appl
28	1609.4	48.8	5558	16	US-10-241-596-137	Sequence 137, App
29	1474	44.7	4700	17	US-10-313-963A-54	Sequence 54, Appl
30	1462.4	44.3	4921	21	US-10-495-491-2	Sequence 2, Appl
31	1462.4	44.3	4935	21	US-10-495-491-1	Sequence 1, Appl
32	1462.4	44.3	4945	21	US-10-495-491-3	Sequence 3, Appl
33	1462.4	44.3	4951	21	US-10-495-491-4	Sequence 4, Appl
34	1417.6	43.0	1511	10	US-09-987-763-30	Sequence 30, Appl
35	1417.6	43.0	5926	13	US-10-024-809-3	Sequence 3, Appl
36	1238.4	37.5	5024	17	US-10-313-963A-52	Sequence 52, Appl
37	1238.4	37.5	5448	21	US-10-671-995A-1	Sequence 2, Appl
38	1238.4	37.5	5454	21	US-10-671-995A-2	Sequence 1, Appl
39	1237.2	37.5	1922	10	US-09-995-988A-32	Sequence 32, Appl
40	1237.2	37.5	1922	17	US-10-420-034A-32	Sequence 32, Appl
41	1237.2	37.5	6270	21	US-10-855-897-2	Sequence 2, Appl
42	1204	36.5	4557	17	US-10-457-372-5	GENERAL INFORMA
43	1204	36.5	6664	18	US-10-270-176-1	Sequence 1, Appl
44	1204	36.5	6668	18	US-10-270-176-2	Sequence 2, Appl
45	1204	36.5	6742	18	US-10-270-176-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Parial sequence of pMal- λ p3 and zinc finger
; OTHER INFORMATION: protein ZFPap3
US-09-765-555-18

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGTCGTGAATGTGAACACGATGAACGTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAAATTCAGGTCGTGAATGTGAACACGATGAACGTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGGTTTCGCGGTGCTGAACCCAGGCCAGCCAGCTTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGGTTTCGCGGTGCTGAACCCAGGCCAGCCAGCTTTCTCGGAAAA 180
QY 181 CCGGGAAAAAGTGGAAGCGGCGATGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CCGGGAAAAAGTGGAAGCGGCGATGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACCACTGGCGGGCAACAGTCTGTTGCTGATTTGGCGTTGCGCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACCACTGGCGGGCAACAGTCTGTTGCTGATTTGGCGTTGCGCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTTGTCGGGCGATTAATACTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTTGTCGGGCGATTAATACTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTTAGAACGAAGCGGCTCGAAGCCCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTGTGATGTTAGAACGAAGCGGCTCGAAGCCCTGTAAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGATCAACGAGTGCAC 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGATCAACGAGTGCAC 480
QY 481 TTGCTGTGGAGCTGCCTGCATAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAGCTGCCTGCATAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGATCGGGCTGGAAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGATCGGGCTGGAAGCATC 600
QY 601 TGGTGCATTTGGGTCAACAGAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCTCGG 660
Db 601 TGGTGCATTTGGGTCAACAGAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCTCGG 660
QY 661 CCGCTCTGGCTGCGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGCTCTGGCTGCGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTCGCAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTCGCAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCGATGSCGCTGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCGATGSCGCTGGGCGCAA 840
QY 841 TGGCGGCCATTCAGAGTCCGGCTGCGCTGGTGGTGCGGATATCTCGTGTAGTGGGATACG 900
Db 841 TGGCGGCCATTCAGAGTCCGGCTGCGCTGGTGGTGCGGATATCTCGTGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGCTTGCAACTCTTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGCTTGCAACTCTTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGGCCGTCTCACTGGTGAAGAAAAACAACCACTCGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCCGTCTCACTGGTGAAGAAAAACAACCACTCGCGCCCAATA 1080
QY 1081 CGCAAAACCGCCTCTCCCGCGCGCTTGGCCGATTCATTAATGTCAGCTGGCAGCAGGTTT 1140

Db 1081 CGCAAAACCGCCTCTCCCGCGCGCTTGGCCGATTCATTAATGTCAGCTGGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTGCAGCTTATCATGCACTGACCGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTCATGTTTGCAGCTTATCATGCACTGACCGGTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCCATCGGAAGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1261 TCAGGAGCCCATCGGAAGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 TGTGCTCAAGCGGCGCATCTCCGTTCTCGATAATGTTTTTTTTCGCGCGCATCATCAACGTTT 1380
Db 1321 TGTGCTCAAGCGGCGCATCTCCGTTCTCGATAATGTTTTTTTTCGCGCGCATCATCAACGTTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTATAGTGTGTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTATAGTGTGTTTCACGA 1500
QY 1501 GCATTCACCAACGAAGCATAGATTTATGAAAATCGAAGAGGTAACTGGTAAATCTCG 1560
Db 1501 GCATTCACCAACGAAGCATAGATTTATGAAAATCGAAGAGGTAACTGGTAAATCTCG 1560
QY 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAAATCGGTAAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAAATCGGTAAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTTCACCGTTGAGCATCCGGAATAAACTGGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTTCACCGTTGAGCATCCGGAATAAACTGGAAGAGAAATTCGCCACAGTT 1680
QY 1681 GCGGCACTGGCGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740
Db 1681 GCGGCACTGGCGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAAACCGTACCTTCACTGGCGCTGATTTGCTGTGACGGGGGTTTATCGTTCAAG 2040
Db 1981 CTGCAAGAAACCGTACCTTCACTGGCGCTGATTTGCTGTGACGGGGGTTTATCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGCGTGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGCGTGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTTAAACCAACACATGAATGACAGACCGGATTC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTTAAACCAACACATGAATGACAGACCGGATTC 2160
QY 2161 TCCATTCGCAAGAGCTGCCCTTTTAAAGCGGAAACGAGCATGACCATCAACCGCCCGTGG 2220

Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGGTGG 2220
Qy 2221 GCATGGTCAACATCGACACACGAGAAAGTGAATTATGTGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGAGAAAGTGAATTATGTGTAAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTCTGCGGTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGCGGTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGCGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAAACAAAGAGCTGCGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
Qy 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGAAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGAAAG 2460
Qy 2461 TTGGCGAAGATCCACGATATTGCCGCCACCATGGAAACGCCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGATATTGCCGCCACCATGGAAACGCCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAAACATCCCGCAGATGTCGCTTCTTGGTATGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAAACATCCCGCAGATGTCGCTTCTTGGTATGCTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGTCAGACTGTGATGAAGCCCTGAAAGACGGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGTCAGACTGTGATGAAGCCCTGAAAGACGGCAGACTAAATTCGAGTCTG 2640
Qy 2641 AACACACACAAATAACAATAACAACTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
Db 2641 AACACACACAAATAACAATAACAACTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
Qy 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAAGCCCTATGCTGTG 2760
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAAGCCCTATGCTGTG 2760
Qy 2761 CCGGAATGTGTAAGTCTTTCAGCAGAGCAGCTCCCTGGTGGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTTCAGCAGAGCAGCTCCCTGGTGGCCACACAGCGTACCCAC 2820
Qy 2821 ACGGGTGAACAAACCGTATAAATGCGCAGAGTGGCGAAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGGTGAACAAACCGTATAAATGCGCAGAGTGGCGAAATCTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATCTGGCGAGAGGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATCTGGCGAGAGGCCATACAAATGTCCAGAAATGTGGC 2940
Qy 2941 AAGTCTTTTCAGCCAGTCCAGCAACCTGTGTGCGCCACCAACGTAATCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTTCAGCCAGTCCAGCAACCTGTGTGCGCCACCAACGTAATCACACCGGGGAGAAG 3000
Qy 3001 CCTATGCTTGTTCGGAAATGTGGTAAGTCTTTCAGCAACAGTGGCTCTTGGTTAGACAC 3060
Db 3001 CCTATGCTTGTTCGGAAATGTGGTAAGTCTTTCAGCAACAGTGGCTCTTGGTTAGACAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAACAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAACAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTAGC 3120
Qy 3121 CAGCGCGCCACCTTGGAACGCCATCAACGCACTCATCTACTGCGGAGAGGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTTGGAACGCCATCAACGCACTCATCTACTGCGGAGAGGCCATACAAATGT 3180
Qy 3181 CCAGAAATGTGGAAGTCTTTTCTCAACTTCAGGCAACTTGGTCCGTCACCAACGTAATCTAC 3240
Db 3181 CCAGAAATGTGGAAGTCTTTTCTCAACTTCAGGCAACTTGGTCCGTCACCAACGTAATCTAC 3240
Qy 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGAGTACCCGTAAGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGAGTACCCGTAAGAGTTCGGGACTACGCT 3300

RESULT 2
US-09-765-555-17
; Sequence 17, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein ZFPm4
US-09-765-555-17

Query Match 98.0%; Score 3232.8; DB 10; Length 3300;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 3258; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGGCATGATAGGCCCGGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGGCATGATAGGCCCGGGAAGAGA 60
Qy 61 GTCAAATTCAGGGTGGTGAATGTGAAACCAAGTAAAGTTATACGATGTCGACAGTATGCGG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAAACCAAGTAAAGTTATACGATGTCGACAGTATGCGG 120
Qy 121 GTGTCCTTATCAGACCCGTTTCCCGCGTGTGAAACCAAGCGCCAGCCACCGTTTCTCGAAAA 180
Db 121 GTGTCCTTATCAGACCCGTTTCCCGCGTGTGAAACCAAGCGCCAGCCACCGTTTCTCGAAAA 180
Qy 181 CCGGAAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CCGGAAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Qy 241 AACAACTGCGGGCAACAGTCGTTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGCAACAGTCGTTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGGCCGTTCGCAAAATTCGCGGGGATTAATCTCGCGCCGATCAACTGCGGTGCGCAGCG 360
Db 301 ACGGCCGTTCGCAAAATTCGCGGGGATTAATCTCGCGCCGATCAACTGCGGTGCGCAGCG 360
Qy 361 TGGTGGTTCGATGTTAGAACGAAAGCGCGCTGAAAGCCTGTAAAGCGCGGTGCAAAATC 420
Db 361 TGGTGGTTCGATGTTAGAACGAAAGCGCGCTGAAAGCCTGTAAAGCGCGGTGCAAAATC 420
Qy 421 TTCTCGCAACCGCTCAGTGGGCTGATCACTAACTATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCAACCGCTCAGTGGGCTGATCACTAACTATCCGCTGGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGAAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTGACCCAGA 540
Db 481 TTGCTGTGAAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGTAACGCACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGTAACGCACTGGCGGTGGAGCATC 600
Qy 601 TGGTCCGATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCCATTTAAGTTCTGTCTCG 660

Db 601 TGGTCGCATTGGGTCACCGAGCAATCGCGCTGTGTAGCGGCCCATTTAAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGGCTCTGGCTGGCTGATATAATCTCACTCGCATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTCTGGCTGGCTGATATAATCTCACTCGCATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCTTCCCACTGGATGCTGGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Db 781 ATGAGGGCATCTTCCCACTGGATGCTGGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Qy 841 TGGCGGCCATTACCGAGTCCGGCTCGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
Db 841 TGGCGGCCATTACCGAGTCCGGCTCGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
Qy 901 AGGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAAAACAGGATTTTC 960
Db 901 AGGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACACGAGCTGGACCGCTTGCCTGCAACTCTCTCAGGCCACGGCGTGA 1020
Db 961 GCCTGCTGGGGCAACACGAGCTGGACCGCTTGCCTGCAACTCTCTCAGGCCACGGCGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGCGACGACAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGCGACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGCTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGCAGCTTATCATCGACTGCAGCGTGACCAATCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGCAGCTTATCATCGACTGCAGCGTGACCAATCTTCTGGCG 1260
Qy 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGCTGTAATCATCTGCATAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGCTGTAATCATCTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTTCGCGGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTTCGCGGACATCATACCGGTT 1380
Qy 1381 CTGGCAATATCTGGAATGAGCTGTTGAACAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAATATCTGGAATGAGCTGTTGAACAATTAATCATCGGCTCGTATAATGTGGA 1440
Qy 1441 ATTGTAGCGGATAACCAATTTTCAACAGGAACAGCCAGTTCGGTTTGGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACCAATTTTCAACAGGAACAGCCAGTTCGGTTTGGTTTTCACGA 1500
Qy 1501 GCACTTCAACCAAGGACCATAGATTATGAAGTGAAGGTAACCTGGTAATCTGG 1560
Db 1501 GCACTTCAACCAAGGACCATAGATTATGAAGTGAAGGTAACCTGGTAATCTGG 1560
Qy 1561 ATTAACCGGATTAAGGCTATACGGTCTCGCTGAAGTTCGGTAAGAAATTCGGAAGAT 1620
Db 1561 ATTAACCGGATTAAGGCTATACGGTCTCGCTGAAGTTCGGTAAGAAATTCGGAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAAATTCACAGGTT 1680
Qy 1681 GCGCACTGGCGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Db 1741 GCGCACTGGCGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740

Db 1681 GCGCACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Qy 1801 CCGTTTACCTCGGATGCGGTACGTTCAACGGCAAGCTGATTGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTCGGATGCGGTACGTTCAACGGCAAGCTGATTGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAACCTCGGAA 1920
Qy 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATTAACGCTGCGCGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATTAACGCTGCGCGAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTTAAACAAACATGATGAGACACCGATTTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTTAAACAAACATGATGAGACACCGATTTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACCGACCAAGTAAATTTGCTGCTGACTGATGAAGTCTG 2280
Db 2221 GCATGCTCAACATCGACACCGACCAAGTAAATTTGCTGCTGACTGATGAAGTCTG 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTGGCTGCTGAGCGCAGGTTATTAACGCCCGCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGGCTGCTGAGCGCAGGTTATTAACGCCCGCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGATTTCTCGAAACCTATCTGCTGACTGATGAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGATTTCTCGAAACCTATCTGCTGACTGATGAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGCTCGCTTCTGCTGATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGCTCGCTTCTGCTGATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Qy 2581 GCGAGCGTCTGACACTGCTGATGAAGCCCTCGAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCGAGCGTCTGACACTGCTGATGAAGCCCTCGAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2700
Db 2641 AACAAACAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2700
Qy 2701 GATCTCTTCTCTGTCGCGCGCGCTCTGAGCGCGCGGGAAGAGCCCTATGCTTGT 2760
Db 2701 GATCTCTTCTCTGTCGCGCGCGCTCTGAGCGCGCGGGAAGAGCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTTTCAGCGCAGAGCTCCCTGCTGGTGGCGCCACCGAGGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTTCAGCGCAGAGCTCCCTGCTGGTGGCGCCACCGAGGTACCCAC 2820

1201 GCNCAATTTCTCATGTTTGAAGCTTATCATGCACTGCAAGGTGCAACCAATGCTTCTGGCG 1260
1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAAGTGTGTAATCACTGCATATATCG 1320
1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAAGTGTGTAATCACTGCATATATCG 1320
1321 TGTGCGTCAAGGGGCACTCCGTTCTGGATTAATGTTTTTGGCGGCAATCAACGGTT 1380
1321 TGTGCGTCAAGGGGCACTCCGTTCTGGATTAATGTTTTTGGCGGCAATCAACGGTT 1380
1381 CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
1381 CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
1441 ATTGTGAGCGGATAACAATTTTACACAGGAACAGCCAGTCCGTTTAGGTGTTTACGA 1500
1441 ATTGTGAGCGGATAACAATTTTACACAGGAACAGCCAGTCCGTTTAGGTGTTTACGA 1500
1501 GCACCTTCAACCAAGGACCATAGATTATGAATACTGAAGAGGTAACTCTGTAATCTGG 1560
1501 GCACCTTCAACCAAGGACCATAGATTATGAATACTGAAGAGGTAACTCTGTAATCTGG 1560
1561 ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
1561 ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGAGAGAAATTCGAGAGTT 1680
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGAGAGAAATTCGAGAGTT 1680
1681 GGGCAATCGCGATGCGCTGACATTAATCTTCTGGGACACAGCCGCTTTGGTGGCTAC 1740
1681 GGGCAATCGCGATGCGCTGACATTAATCTTCTGGGACACAGCCGCTTTGGTGGCTAC 1740
1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGACAAAGGTTTCCAGGACAAAGCTGAT 1800
1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGACAAAGGTTTCCAGGACAAAGCTGAT 1800
1801 CCGTTTACCTGGATGCGCTGATTAACAAACCGGACAAAGTGTATCCCGATCGCTGTT 1860
1801 CCGTTTACCTGGATGCGCTGATTAACAAACCGGACAAAGTGTATCCCGATCGCTGTT 1860
1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTGCGAACCCTGCAAAACCTGGGAA 1920
1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTGCGAACCCTGCAAAACCTGGGAA 1920
1921 GAGATCCGCGCTGGATAAAGAACTGAAGAGGAAAGGTAAGAGCGGCTGATGTTCAAC 1980
1921 GAGATCCGCGCTGGATAAAGAACTGAAGAGGAAAGGTAAGAGCGGCTGATGTTCAAC 1980
1981 CTGCAAGAACCTGATCTTCACTGCGCGCTGATGCTGTAAGCGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCTGATCTTCACTGCGCGCTGATGCTGTAAGCGGGGTTATGCGTTCAAG 2040
2041 TATGAAACGCGCAAGTACGACATTAAGAGCTGCGGCTGGATTAAGCTGCGGCAAGCG 2100
2041 TATGAAACGCGCAAGTACGACATTAAGAGCTGCGGCTGGATTAAGCTGCGGCAAGCG 2100
2101 GGTCTGACCTTCTCGTTGATCTGATTAATAAACAACACATGATGAGACACCGATTAC 2160
2101 GGTCTGACCTTCTCGTTGATCTGATTAATAAACAACACATGATGAGACACCGATTAC 2160
2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAACAAGGATGACCATCAACCGGCGCTGG 2220
2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAACAAGGATGACCATCAACCGGCGCTGG 2220
2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGGACCTTC 2280
2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGGACCTTC 2280
2281 AAGGGTCAACCATCCAAACCGTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGGCGAGT 2340
2281 AAGGGTCAACCATCCAAACCGTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGGCGAGT 2340

RESULT 4

US-09-765-555-15
; Sequence 15, Application US/09765555
; Publication No. US2003003735A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate

2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAGAG 2460
2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAGAG 2460
2461 TTGGCGAAAGATCCACGATATTGCGCCACCATGGAACACGCCAAGAGGTGAATCATG 2520
2461 TTGGCGAAAGATCCACGATATTGCGCCACCATGGAACACGCCAAGAGGTGAATCATG 2520
2521 CCGAAACATCCGCGAGATGCTGGCTTTCTGCTGATGCGGTGCTGCTGCGGTGATCAACGCC 2580
2521 CCGAAACATCCGCGAGATGCTGGCTTTCTGCTGATGCGGTGCTGCTGCGGTGATCAACGCC 2580
2581 GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACCGCAGAGCTAAATTCGAGCTCG 2640
2581 GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACCGCAGAGCTAAATTCGAGCTCG 2640
2641 AACAAACAAACAAATAAATAAACAACCTCGGGATCGAGGAAGGATTTTCAAGATTC 2700
2641 AACAAACAAACAAATAAATAAACAACCTCGGGATCGAGGAAGGATTTTCAAGATTC 2700
2701 GGATCTCTTCTCTGTCGCGCCAGGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
2701 GGATCTCTTCTCTGTCGCGCCAGGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
2761 CCGGAATGTGTTAAGTCTTTTACGACAGAGCAGCTCCCTGCTGCGCCACACAGCGTACCCAC 2820
2761 CCGGAATGTGTTAAGTCTTTTCTCTCAGAGCTCTCACTGCTGCGCCACACAGCGTACCCAC 2820
2821 ACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGGCAAAATCTTTTACGCAAGTCCAGCAAC 2880
2821 ACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGGCAAAATCTTTTACGCAAGTCCAGCAAC 2880
2881 CTGCTGCGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
2881 CTGCTGCGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
2941 AAGTCTTTTACGCGATCCAGCAACCTGCTGCGCCACCAACGTACTCACACCGGGGAGAAG 3000
2941 AAGTCTTTTCTCGGTCTGACAACTCTGCTGCGGACCAACGTACTCACACCGGGGAGAAG 3000
3001 CCTATGCTTGTCCGGAATGTGGTAACTCTTTCAGCACAGTGGCTCTTGTGTTAGCAC 3060
3001 CCTATGCTTGTCCGGAATGTGGTAACTCTTTCAGCACAGTGGCTCTTGTGTTAGCAC 3060
3061 CAGCGTACCCACACCGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGGCAAACTCTTTTAGC 3120
3061 CAGCGTACCCACACCGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGGCAAACTCTTTTAGC 3120
3121 CAGCGCGCCACCTGGAACCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
3121 CAGCGCGCCACCTGCGGAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
3181 CCAGAATGTGGCAAGTCTTTTCAACTTTCAGCACTTCTGCTGCGTCCGTCACCAACGTA 3240
3181 CCAGAATGTGGCAAGTCTTTTCAACTTTCAGCACTTCTGCTGCGTCCGTCACCAACGTA 3240
3241 ACCGGTAAAAAACTAGTGGCGCGGCGGCGGAGTACCCGTAGCAGCTTCCGGACTACGCT 3300
3241 ACCGGTAAAAAACTAGTGGCGCGGCGGCGGAGTACCCGTAGCAGCTTCCGGACTACGCT 3300

; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger
; OTHER INFORMATION: protein 2FPm2
us-09-765-555-15

Query Match 97.8%; Score 3226.4; DB 10; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 CGGACACCATCGAATGGTGCAGAACTTTTCGGGGTATGCGCATGATAGCGCCCGGAGAGA 60
DB 1 CGGACACCATCGAATGGTGCAGAACTTTTCGGGGTATGCGCATGATAGCGCCCGGAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACTGATACCATGTCGACAGTATGCGG 120
DB 61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACTGATACCATGTCGACAGTATGCGG 120
QY 121 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAA 180
DB 121 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAA 180
QY 181 CGCGGGAAAACTGGAAGCGGATGGGGAGCTGAATTTACATTTCCCAACCCGCTGGCAC 240
DB 181 CGCGGGAAAACTGGAAGCGGATGGGGAGCTGAATTTACATTTCCCAACCCGCTGGCAC 240
QY 241 AACAACTGCGGGGAAAAAGTGTGCTGATTTGGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGCGGGGAAAAAGTGTGCTGATTTGGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTFCGAAATTTGCGGGGATTAATCTCGGCCGATCAACTGGGTGCGACG 360
DB 301 ACGCGCGTFCGAAATTTGCGGGGATTAATCTCGGCCGATCAACTGGGTGCGACG 360
QY 361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGGTGCACAAATC 420
DB 361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACACAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACACAGA 540
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACCGGCTGGGCGGTGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACCGGCTGGGCGGTGAGCATC 600
QY 601 TGGTTCGATTTGGTTCACAGCAAACTCGGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG 660
DB 601 TGGTTCGATTTGGTTCACAGCAAACTCGGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG 660
QY 661 CGCGTCTGGCTCTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGGCTCTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGCATGCGATGCGATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780

DB 721 CGGAACGGGAAGCGCATGCGATGCGATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCCATCGCATGCTGGTTCGCAACAGATCAGATGCGCTGGCGCGAA 840
DB 781 ATGAGGGCATCGTTCCTCCATCGCATGCTGGTTCGCAACAGATCAGATGCGCTGGCGCGAA 840
QY 841 TGGCGGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATAG 900
DB 841 TGGCGGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGCTTAAACACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGCTTAAACACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCCACGCGGTGA 1020
DB 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCCACGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGGCGCTCTCAGTGTGAAAGAAACACACCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGGCGCTCTCAGTGTGAAAGAAACACACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGTGGCAGCAGCAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGTGGCAGCAGCAGGTTT 1140
QY 1141 CCAGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTATTAG 1200
DB 1141 CCAGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTATTAG 1200
QY 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGCAAGTGCACCAATGTTCTGGCG 1260
DB 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGCAAGTGCACCAATGTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGCTGTCAGGTCGTAATCACTGATTAATTCG 1320
DB 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGCTGTCAGGTCGTAATCACTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
DB 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
DB 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATACAATTTTACACAGCAACACGAGTCGCTTGTAGTGTTCACGA 1500
DB 1441 ATTGTAGCGGATACAATTTTACACAGCAACACGAGTCGCTTGTAGTGTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAACCTGTTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAACCTGTTAATCTGG 1560
QY 1561 ATTAAACGGGATAAAGGCTATAACCGCTCTCGCTGAGTTCGGTAAAGAAATTCAGAAAGAT 1620
DB 1561 ATTAAACGGGATAAAGGCTATAACCGCTCTCGCTGAGTTCGGTAAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACACGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACACGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTTGGCTGGAATCAACCCGGAACAAAGCGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGCTGTTGGCTGGAATCAACCCGGAACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACAGTTCACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
DB 1801 CCGTTTACCTGGATGCGGTACAGTTCACGCAAGCTGATTGCTTACCCGATCGCTGT 1860

QY 1861 GAAGGTTATCGCTGATTTTATAACAAAGATCTGTCGCAACCCGCCAAAACCTGGGAA 1920
DB |||||
QY 1861 GAAGGTTATCGCTGATTTTATAACAAAGATCTGTCGCAACCCGCCAAAACCTGGGAA 1920
DB |||||
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGAGTAAAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGAGTAAAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
DB |||||
QY 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
DB |||||
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2101 GGTCTGACCTTCTGCTGACCTGATTAAGAACTGAAGCGTGGCGTGGATAACGCTGGCGGAAAGCG 2160
DB |||||
QY 2101 GGTCTGACCTTCTGCTGACCTGATTAAGAACTGAAGCGTGGCGTGGATAACGCTGGCGGAAAGCG 2160
DB |||||
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGCGCAACAGCGATGACCATCAACGGCCGCTGG 2220
DB |||||
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGCGCAACAGCGATGACCATCAACGGCCGCTGG 2220
DB |||||
QY 2221 GCATGGTCAACATCGACACGCAAGAGTGAATTATGTTGTAACGGTACTGCGGACCTTC 2280
DB |||||
QY 2221 GCATGGTCAACATCGACACGCAAGAGTGAATTATGTTGTAACGGTACTGCGGACCTTC 2280
DB |||||
QY 2281 AAGGTCACCAACCTCAACCGTTCGCTGCTGCTGAGCGGAGGTATTAAGCGCGCCAGT 2340
DB |||||
QY 2281 AAGGTCACCAACCTCAACCGTTCGCTGCTGCTGAGCGGAGGTATTAAGCGCGCCAGT 2340
DB |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
DB |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
DB |||||
QY 2401 GAAGCGGTTAATGAACAAACCGTGGTGGCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
DB |||||
QY 2401 GAAGCGGTTAATGAACAAACCGTGGTGGCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
DB |||||
QY 2461 TTGGCAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2461 TTGGCAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGATCAAGGCC 2580
DB |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGATCAAGGCC 2580
DB |||||
QY 2581 GCCAGCGTCTGACGCTGATGAGCCCTGGAAGACGCGGAGACTAATTCGAGCTCG 2640
DB |||||
QY 2581 GCCAGCGTCTGACGCTGATGAGCCCTGGAAGACGCGGAGACTAATTCGAGCTCG 2640
DB |||||
QY 2641 AACCAACAAACAAATAACAAACAACTCGGGATCGAGGGAAGATTTTCAGAAATTC 2700
DB |||||
QY 2641 AACCAACAAACAAATAACAAACAACTCGGGATCGAGGGAAGATTTTCAGAAATTC 2700
DB |||||
QY 2701 GGATCTCTTCTCTGTCGCGCGGCTCTGAGCGCGGGGAGAGCCCTATGCTTGT 2760
DB |||||
QY 2701 GGATCTCTTCTCTGTCGCGCGGCTCTGAGCGCGGGGAGAGCCCTATGCTTGT 2760
DB |||||
QY 2761 CCGGAATGTGGTAAAGTCTCTTCAGCGCAGAGCTCTGCTGCTGCGGAGAGCCCTATGCTTGT 2820
DB |||||
QY 2761 CCGGAATGTGGTAAAGTCTCTTCAGCGCTCTCCTGCTGCTGCGGAGAGCCCTATGCTTGT 2820
DB |||||
QY 2821 ACGGGTGAAAAACCGTATAATGCCAGAGTGGGCAAACTCTTTTACGCAAGTCCAGCAAC 2880
DB |||||
QY 2821 ACGGGTGAAAAACCGTATAATGCCAGAGTGGGCAAACTCTTTTACGCAAGTCCAGCAAC 2880
DB |||||
QY 2881 CTGGTGGCCATCAACGCACTTACTGCGGAGAGCCATCAAAATGTCAGAAATGTGGC 2940
DB |||||
QY 2881 CTGGTGGCCATCAACGCACTTACTGCGGAGAGCCATCAAAATGTCAGAAATGTGGC 2940
DB |||||

QY 2941 AAGTCTTTACGCCAGTCCAGCAACCTGCTGCGCGCCACCAACGTACTCACACCGGGAGAG 3000
DB |||||
QY 2941 AAGTCTTTTCTCTCGGTCTGCAATCTCGTCCGGCACCAACGTACTCACACCGGGAGAG 3000
DB |||||
QY 3001 CCCTATGCTTCTCGCGAATGTGGTAAAGTCTTTCAGCACCAAGTCTCTCTGCTTGTAGACAC 3060
DB |||||
QY 3001 CCCTATGCTTCTCGCGAATGTGGTAAAGTCTTTCAGCGCGCAGGATACCTGCTGCGCAC 3060
DB |||||
QY 3061 CAGGTTACCCACACGCGGTGAAAAACCGTATAAAATGCCAGAGTCCGCAAAATCTTTTACG 3120
DB |||||
QY 3061 CAGGTTACCCACACGCGGTGAAAAACCGTATAAAATGCCAGAGTCCGCAAAATCTTTTACG 3120
DB |||||
QY 3121 CAGCGCGCCACCTGGAAAGCCATCAACGCACTCATACTGCGGAGAGCCATACAAATGT 3180
DB |||||
QY 3121 CAGCGCGCCACCTGGCGGCAATCAACGCACTCATACTGCGGAGAGCCATACAAATGT 3180
DB |||||
QY 3181 CCAGAATGTGGCAAGTCTTTCTCAACTTCAGGCAACTTGGTCCGTCACCAACGTACTCAC 3240
DB |||||
QY 3181 CCAGAATGTGGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCAACGTACTCAC 3240
DB |||||
QY 3241 ACCGTTAAAAAACTAGTGGCCAGCGCCGCGCAGTACCCGTACGACGTTCCGGACTAGCT 3300
DB |||||
QY 3241 ACCGTTAAAAAACTAGTGGCCAGCGCCGCGCAGTACCCGTACGACGTTCCGGACTAGCT 3300
DB |||||

RESULT 5
US-09-765-555-16
; Sequence 16, Application US/09765555
; Publication No. US2003003735A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-28014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m3 and zinc finger
; OTHER INFORMATION: protein ZFPm3
US-09-765-555-16

Query Match 97.5%; Score 3216.8; DB 10; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGCGTATGCGCATGATAGCGCCCGGAAGAGA 60
DB |||||
QY 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGCGTATGCGCATGATAGCGCCCGGAAGAGA 60
DB |||||
QY 61 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTGATACGATGTCGAGAGTATGCCG 120
DB |||||
QY 61 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTGATACGATGTCGAGAGTATGCCG 120
DB |||||
QY 121 GTGCTCTTTATCAGACCGCTTTCCCGGTGTAATGTGAACCAAGTAACTGATACGATGTCGAGAGTATGCCG 180
DB |||||
QY 121 GTGCTCTTTATCAGACCGCTTTCCCGGTGTAATGTGAACCAAGTAACTGATACGATGTCGAGAGTATGCCG 180
DB |||||
QY 181 CCGGGGAAAAAGTGAAGCGCGGATGGCGAGTGAATTTACATTTCCAAACCGCGTGGCAC 240
DB |||||
QY 181 CCGGGGAAAAAGTGAAGCGCGGATGGCGAGTGAATTTACATTTCCAAACCGCGTGGCAC 240
DB |||||
QY 241 AACCACTGCGGCAACCAAGTGTGCTGATTTGGGTGGCACTCCAGTCTGGCCCTGC 300
DB |||||

Db 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGGCGCTCGCAAAATTGTTCGGCGGATTAATATCTCGCGCGGATCAATCTGGGTGCCAGC 360
Db 301 ACGGCGCTCGCAAAATTGTTCGGCGGATTAATATCTCGCGCGGATCAATCTGGGTGCCAGC 360
Qy 361 TGGTGGTCTGCATGTAGAACGAAAGCGGCTCGAAGCTGTAAAGCGGCTGACACAATC 420
Db 361 TGGTGGTCTGCATGTAGAACGAAAGCGGCTCGAAGCTGTAAAGCGGCTGACACAATC 420
Qy 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTGC 480
Db 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTGC 480
Qy 481 TTGCTGTGGAGCTGCCCTGACATAATGTTCCGGGTTATTTCTTGGATGCTCTGACAC 540
Db 481 TTGCTGTGGAGCTGCCCTGACATAATGTTCCGGGTTATTTCTTGGATGCTCTGACAC 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCACTGGGCGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGGTCAACAGCAAAATCGGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCG 660
Db 601 TGGTGGCATTTGGGTCAACAGCAAAATCGGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCG 660
Qy 661 CGGCTCTGGCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGGCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGCA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGCA 840
Qy 841 TGGCGCGCATTAACGAGTCCGGGCTGCGGTTGTCGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGCGCATTAACGAGTCCGGGCTGCGGTTGTCGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCAATCAACAGGATTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCAATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACAGCGTGAACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGAACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTA 1020
Qy 1021 AGGGCAATCAGTGTGGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGTGTGGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 GCAAAACGCTCTCCCGCGGTTGGCGGATTCATTAAATGCAAGTGGCAACAGGTTT 1140
Db 1081 GCAAAACGCTCTCCCGCGGTTGGCGGATTCATTAAATGCAAGTGGCAACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGTGGCAACCAATGCTTCGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGTGGCAACCAATGCTTCGGCG 1260
Qy 1261 TCAGGAGCATTCGGAAGCTGTGGTATGCGCTGTGAGGTCTGTAATCATCTGCATAAATTCG 1320
Db 1261 TCAGGAGCATTCGGAAGCTGTGGTATGCGCTGTGAGGTCTGTAATCATCTGCATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380

Qy 1381 CTGCAAAATATTTCTGAATGAGCTGTTGAATAATTAATCATCGGCTCGTATATGTTGCGA 1440
Db 1381 CTGCAAAATATTTCTGAATGAGCTGTTGAATAATTAATCATCGGCTCGTATATGTTGCGA 1440
Qy 1441 ATTCTGAGCGGATAAACAATTTTACACAGGAAACACGAGTCCGTTTAGTGTGTTTTCAGA 1500
Db 1441 ATTCTGAGCGGATAAACAATTTTACACAGGAAACACGAGTCCGTTTAGTGTGTTTTCAGA 1500
Qy 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTCGTAACTCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTCGTAACTCTGG 1560
Qy 1561 ATTTAACGGCGATAAGGCTATAACGGTCTCGCTGAGTCCGTTAGAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAACGGCGATAAGGCTATAACGGTCTCGCTGAGTCCGTTAGAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTTAAAGTCAACGGTTGAGCATCCGATATAACTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCAACGGTTGAGCATCCGATATAACTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTGTTGGCTGAAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Qy 1801 CCGTTTACCTGGGATGCGTACGTTTACAACGGCAAGCTGATTGCTTACCCTCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTTACAACGGCAAGCTGATTGCTTACCCTCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGCGCTGATTCGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGCGCTGATTCGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACCGTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACCGTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGTGCCTTTAATAAGGCGAAACGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGTGCCTTTAATAAGGCGAAACGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACACGAGCAAAAGTGAATTAATGGTGAACGCTGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACACGAGCAAAAGTGAATTAATGGTGAACGCTGCGACCTTC 2280
Qy 2281 AAGGTCACCAATCCAAACCGTTGCGCTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTGCGCTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGCTCG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTTAGCGCTGAAAGTCTTACGAGGAAG 2460

```
QY 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATGGAAAGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATGGAAAGCGCCAGAAAGGTGAATCATG 2520
QY 2521 CGGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CGGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGCTCAGATGTCGATGAAGCCCTTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTCAGATGTCGATGAAGCCCTTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACACAAACAATAACAATAACAACAACTTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
Db 2641 AACACACAAACAATAACAATAACAACAACTTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTCTGTCGCGCCAGCGGCCCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCCTCTTCTCTGTCGCGCCAGCGGCCCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTCTTCAGCCAGAGCAGCTCCCTGGTGGCCACCCAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTCTTCAGCGATCCTTGCCACCTGTTGCGCCACCGAGTACCCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAACTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAACTTTTAGCCAGCGGGCTCC 2880
QY 2881 CTGGTGGCCCATCAACGACATCATCTGCGGAGAGCCATACAATGTCGAGATGTTGGC 2940
Db 2881 CTGGTGGCCCATCAACGACATCATCTGCGGAGAGCCATACAATGTCGAGATGTTGGC 2940
QY 2941 AAGTCTTTACGCCAGTCAGCAACCTGTTGGTGGCCACCAACGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTACGCCAGTCAGCAACCTGTTGGTGGCCACCAACGTACTCACACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTCTCGGAGTGTGAAGTCTTACGACCAAGTGGCTCTGTTAGTACAC 3060
Db 3001 CCCTATGCTTCTCGGAGTGTGAAGTCTTACGACCAAGTGGCTCTGTTAGTACAC 3060
QY 3061 CAGCGTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAACTTTTAGC 3120
Db 3061 CAGCGTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGGCAAACTTTTAGT 3120
QY 3121 CAGCGCGCCACCTGGGAACGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 GACTGCGCGCACTTGTCTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCAACTCAGGCACTTGGTCCGTACCAACGTAAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCAACTCAGGCACTTGGTCCGTACCAACGTAAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCGAGCGCGGCCAGTACCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCGAGCGCGGCCAGTACCGTACGACGTTCCGGACTACGCT 3300

RESULT 6
US-10-257-384A-1
; Sequence 1, Application US/10257384A
; Publication No. US20040087524A1
; GENERAL INFORMATION:
; APPLICANT: Wiederanders, Bernd
; TITLE OF INVENTION: Maubach, Gunter
; FILE OF INVENTION: Agent for postoperative use after removal of bone tumors
; FILE REFERENCE: 2945-101
; CURRENT APPLICATION NUMBER: US/10/257,384A
; PRIOR FILING DATE: 2003-08-21
; PRIOR FILING DATE: 2001-04-18
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 19
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7373
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Expression vector pMalc2cchbnp2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)..(3453)
; OTHER INFORMATION: /product = fusion protein MBP Cystatin C BMP-2
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2689)
; OTHER INFORMATION: Faktor Xa Protease /position=388
; OTHER INFORMATION: (Amino acid sequence)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3067)..(3108)
; OTHER INFORMATION: /spacer between Cystatin C andBMP-2
; PUBLICATION INFORMATION:
; AUTHORS: Zwick, M B.
; TITLE: Expression vector pMal-X, complete sequence
; JOURNAL: Anal. Biochem.
; VOLUME: 264
; ISSUE: 1
; PAGES: 87-97
; DATE: 1998-11-01
; DATABASE ACCESSION NUMBER: Entrez Nucleotide database/AF031813
; DATABASE ENTRY DATE: 2001-05-07
; RELEVANT RESIDUES: 1528 - 3453
US-10-257-384A-1
```

```
Query Match 82.0%; Score 2705.8; DB 18; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCGCGGTATGCATGATAGCCGCCGAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCGCGGTATGCATGATAGCCGCCGAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTG AAAACCAAGTAACTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTG AAAACCAAGTAACTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCCAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCCAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACCACTGCGGGCAAAACAGTGGTGTCTGATTTGGGGTTCGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACCACTGCGGGCAAAACAGTGGTGTCTGATTTGGGGTTCGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGGCCCATCACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGGCCCATCACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGGTAGAAACGAAGCGCGTCCGAAGCCCTGTAAAGCGGGGTGCAAAATC 420
Db 361 TGGTGGTGTGATGGTAGAAACGAAGCGCGTCCGAAGCCCTGTAAAGCGGGGTGCAAAATC 420
QY 421 TTCTCGCGCAACCGGTGAGTGGGTGATCATTAATATCCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGGTGAGTGGGTGATCATTAATATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGCGGTTATTTCTGTATGTCCTGACCCA 540
Db 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGCGGTTATTTCTGTATGTCCTGACCCA 540
```

QY 541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCACTGGCGGTGGAGCATC 600
QY 601 TGGTCCGATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCCGATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATTAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATTAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAAACGGAGGCGACTGGAGTGCATGTCGGTTCCTGCGTTCCTGCGTTCCTGCGTTCCTGCG 780
Db 721 CGGAAACGGAGGCGACTGGAGTGCATGTCGGTTCCTGCGTTCCTGCGTTCCTGCGTTCCTGCG 780
QY 781 ATGAGGGATCTTCCCACTGCGATGTCGGTTCGCAACGATCAGATGGCGCTGGCGCGCAA 840
Db 781 ATGAGGGATCTTCCCACTGCGATGTCGGTTCGCAACGATCAGATGGCGCTGGCGCGCAA 840
QY 841 TGGCGGCCATTAACGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTAACGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACGGAACAGCTCATGTTATATCCCGCGTTAAACACCATCAAAAGGATTTTC 960
Db 901 ACGATACGGAACAGCTCATGTTATATCCCGCGTTAAACACCATCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCACTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCACTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACGCTCTCCCGCGGTTGGCGGATTAATGAGCTGGCAGCAGAGGTTT 1140
Db 1081 CGCAAAACGCTCTCCCGCGGTTGGCGGATTAATGAGCTGGCAGCAGAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTAGTTAGTCTCACTATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTAGTTAGTCTCACTATTAG 1200
QY 1201 GCACAAATCTCATTTGTGACGCTTATCATGACCTGCAGCGTGCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATTTGTGACGCTTATCATGACCTGCAGCGTGCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGCTCGTAATCACTGCATTAATCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGCTCGTAATCACTGCATTAATCG 1320
QY 1321 TGTGCTCAAGCGCACTCCGCTCTGGATAATGTTTTTGGCGGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCGCTCTGGATAATGTTTTTGGCGGACATCAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTGGA 1440
QY 1441 ATTGTAGCGGATTAACAAATTCACAGGAACAGCGAGTCCGTTAGGTGTTTTCAGGA 1500
Db 1441 ATTGTAGCGGATTAACAAATTCACAGGAACAGCGAGTCCGTTAGGTGTTTTCAGGA 1500
QY 1501 GCACCTTCAACAAAGGACATAGATTATGAATACTGAAGAGGTAACCTGCTAATCTCG 1560
Db 1501 GCACCTTCAACAAAGGACATAGATTATGAATACTGAAGAGGTAACCTGCTAATCTCG 1560
QY 1561 ATTTAAACGGCGATAAAGGCTTAAACGCTCTCGCTGAAGTCCGTAAGAAATTCAGAGAAAGAT 1620
Db 1561 ATTTAAACGGCGATAAAGGCTTAAACGCTCTCGCTGAAGTCCGTAAGAAATTCAGAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCACAGGTT 1680
QY 1681 CGGCAACTGCGGATGGCCCTTGACATTAATCTTCTGGGCACACGACCCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGGATGGCCCTTGACATTAATCTTCTGGGCACACGACCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGACAAAGCGTTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGACAAAGCGTTTCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCGTACGTTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGGCGTACGTTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTAAAGACGTTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTAAAGACGTTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGATGATGACAGACCCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGATGATGACAGACCCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGGCGAAACAGAGATGACCATCAAGCGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAGGCGAAACAGAGATGACCATCAAGCGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGCTGAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGCTGAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCAACCATCAACACCGTTGCGTGTGAGCGCAGGATTAATAAGCGCCCGAGT 2340
Db 2281 AAGGTCAACCATCAACACCGTTGCGTGTGAGCGCAGGATTAATAAGCGCCCGAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGGCTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCAAAAGATTCACATTAATGCGCCCAACCATGGAAGAAACGCCAGAAAGGTAATCATG 2520
Db 2461 TTGGCAAAAGATTCACATTAATGCGCCCAACCATGGAAGAAACGCCAGAAAGGTAATCATG 2520
QY 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGGTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGGTGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGGCGAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGGCGAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACAACTTCGCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAACAACTTCGCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGCTGGCGCCAGCGCGCTCGAGGCCGG 2741

Db 2701 GGATCCTCAGTCCCGCAAGCCGCGCGCTGGTGGAGG 2741

RESULT 7
US-10-149-472-5
/ Sequence 5, Application US/10149472
/ Publication No. US20040029204A1
/ GENERAL INFORMATION:
/ APPLICANT: GOUBIN-GRAMATICA, FRANCOISE
/ APPLICANT: DUCOMMUN, BERNARD
/ APPLICANT: PREVOST, GREGOIRE
/ TITLE OF INVENTION: METHOD FOR OBTAINING HUMAN CDC25 PHOSPHATASES AND METHOD
/ FILE REFERENCE: bml-427.065
/ CURRENT APPLICATION NUMBER: US/10/149,472
/ PRIOR FILING DATE: 2002-06-07
/ PRIOR FILING DATE: 2000-12-13
/ PRIOR APPLICATION NUMBER: FR 99/15722
/ PRIOR FILING DATE: 1999-12-14
/ PRIOR APPLICATION NUMBER: FR 00/06883
/ PRIOR FILING DATE: 2000-05-30
/ PRIOR APPLICATION NUMBER: FR 00/12008
/ PRIOR FILING DATE: 2000-09-21
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 5
/ LENGTH: 8101
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ OTHER INFORMATION: JM109 / pMAL - Hs Cdc25C strain
US-10-149-472-5

Query Match 82.0%; Score 2705.8; DB 17; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60

Qy 61 GTCGAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTGCGAGAGTATCCG 120
Db 61 GTCGAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTGCGAGAGTATCCG 120

Qy 121 GTGCTCTTATCAGACCGGTTTCCCGCGTGTGAACCAAGCGGCGAGTGTGCGGAA 180
Db 121 GTGCTCTTATCAGACCGGTTTCCCGCGTGTGAACCAAGCGGCGAGTGTGCGGAA 180

Qy 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATTAATTCACCAACCGCGTGGCAC 240
Db 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATTAATTCACCAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGCAACAGTCGTTGCTGATGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCGTTGCTGATGGCGTGGCCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGCGCGTGCAGAAATGTGCGCGGATTAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTGCAGAAATGTGCGCGGATTAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360

Qy 361 TGGTGGTGTGATGGTGAAGCAAGCGCGTGCAGAGCTGTAAAGCGCGGCTGCACAATC 420
Db 361 TGGTGGTGTGATGGTGAAGCAAGCGCGTGCAGAGCTGTAAAGCGCGGCTGCACAATC 420

Qy 421 TTCTCGCGCAACCGGTGAGTGGGTGATCATTAATCTCGCTGGATGACCAAGTCCCA 480
Db 421 TTCTCGCGCAACCGGTGAGTGGGTGATCATTAATCTCGCTGGATGACCAAGTCCCA 480

Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGTTCTTGACAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGTTCTTGACAGA 540

Db 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGTTCTTGACAGA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCACTGGGCGTGGAGCATC 600
Qy 601 TGGTCCGATTTGGGTGACCAAGCAATCGCGCTGTTAGCGGGCCATTAAGTTCCTCTCGG 660
Db 601 TGGTCCGATTTGGGTGACCAAGCAATCGCGCTGTTAGCGGGCCATTAAGTTCCTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACCGGAAGGCGACTGGAAGTCCCATGTCGGTTTTTCAACAAACCATCAAAATGCTGA 780
Db 721 CGGAACCGGAAGGCGACTGGAAGTCCCATGTCGGTTTTTCAACAAACCATCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCGTGGCGCAA 840
Db 781 ATGAGGGCATCGTTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCGTGGCGCAA 840
Qy 841 TGGCGCCATTACCGAGTCCGGGCTGCGCGTGGTGGGATATCTCGTATGCGGATACG 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTGCGCGTGGTGGGATATCTCGTATGCGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAAGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAAGATTTTC 960
Qy 961 GCCTGCTGGGGCAAAACCGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACCGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGTAAGAAAACCAACCTGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGTAAGAAAACCAACCTGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCAAGTGGCAAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCAAGTGGCAAGCAGGTTT 1140
Qy 1141 CCGCACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 1200
Db 1141 CCGCACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 1200
Qy 1201 GCACAAATTCATGTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTCATGTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGCGAGCATCGGAAGCTGTGATGCTGTGCAAGTCGTAATCACTGCAATAATTCG 1320
Db 1261 TCAGGCGAGCATCGGAAGCTGTGATGCTGTGCAAGTCGTAATCACTGCAATAATTCG 1320
Qy 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAACCGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAACCGTT 1380
Qy 1381 CTGGCAAAATTCGAAATGAGCTGTGCAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATTCGAAATGAGCTGTGCAATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGGTGTTTTCCGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGGTGTTTTCCGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAGCTGAAGAGGTAAGTGGTATCTCG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAGCTGAAGAGGTAAGTGGTATCTCG 1560
Qy 1561 ATTAAACGGCGATAAGAGGCTATTAACCGTCTCGCTGAAGTGGTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAAACGGCGATAAGAGGCTATTAACCGTCTCGCTGAAGTGGTAAGAAATTCGAGAAGAT 1620

Qy 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTTCCACAGGTT 1680
Db |||||||
Qy 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTTCCACAGGTT 1680
Db |||||||
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTCGGGCACACGCGCTTTGGTGGCTAC 1740
Db |||||||
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTCGGGCACACGCGCTTTGGTGGCTAC 1740
Db |||||||
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db |||||||
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db |||||||
Qy 1801 CCGTTTACCTGGATGCGGTAGCTTACAACGCGAAGCTGATTGCTTACCCGATCCTGTT 1860
Db |||||||
Qy 1801 CCGTTTACCTGGATGCGGTAGCTTACAACGCGAAGCTGATTGCTTACCCGATCCTGTT 1860
Db |||||||
Qy 1861 GAAAGGTTATCGCTGATTTATAAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db |||||||
Qy 1861 GAAAGGTTATCGCTGATTTATAAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db |||||||
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGTGTATTTCAAC 1980
Db |||||||
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGTGTATTTCAAC 1980
Db |||||||
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGGTGATGCTGTGACGGGGTTATGCGTTCAAG 2040
Db |||||||
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGGTGATGCTGTGACGGGGTTATGCGTTCAAG 2040
Db |||||||
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGGTGGATAAGCTGGCGGAAAGCG 2100
Db |||||||
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGGTGGATAAGCTGGCGGAAAGCG 2100
Db |||||||
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATCAATGAATGACACACCGATTAC 2160
Db |||||||
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATCAATGAATGACACACCGATTAC 2160
Db |||||||
Qy 2161 TCATCGCAGAAAGCTGCTTTAATAAGCGGAAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db |||||||
Qy 2161 TCATCGCAGAAAGCTGCTTTAATAAGCGGAAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db |||||||
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGGACCTTC 2280
Db |||||||
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGGACCTTC 2280
Db |||||||
Qy 2281 AAGGTCACACCATCAAAACCGTTCTGTCGCTGCTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db |||||||
Qy 2281 AAGGTCACACCATCAAAACCGTTCTGTCGCTGCTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db |||||||
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGACTGATGAAGGTCTG 2400
Db |||||||
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGACTGATGAAGGTCTG 2400
Db |||||||
Qy 2401 GAAAGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACAGGAAGAG 2460
Db |||||||
Qy 2401 GAAAGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACAGGAAGAG 2460
Db |||||||
Qy 2461 TTGGCGAAAGATCCACGTATTGCGGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db |||||||
Qy 2461 TTGGCGAAAGATCCACGTATTGCGGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db |||||||
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db |||||||
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db |||||||
Qy 2581 GCCAGGGTCTGACACTGATGAAGCCCTGAAAGAGCGCGACACTAATTCGAGCTCG 2640
Db |||||||
Qy 2581 GCCAGGGTCTGACACTGATGAAGCCCTGAAAGAGCGCGACACTAATTCGAGCTCG 2640
Db |||||||
Qy 2641 AACAAACAACAATAAACAACACCTCGGATCGGAGGAGGATTTTCAGAAATTC 2700
Db |||||||
Qy 2641 AACAAACAACAATAAACAACACCTCGGATCGGAGGAGGATTTTCAGAAATTC 2700
Db |||||||

Qy 2701 GGATCCTCT 2709
Db |||||||
Qy 2701 GGATCCTCT 2709
Db |||||||
RESULT 8
US-10-343-859-8
; Sequence 8, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 6648
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
; OTHER INFORMATION: pMAL-c2x
US-10-343-859-8
Query Match 81.8%; Score 2701; DB 19; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60
Db |||||||
Qy 1 CCGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60
Db |||||||
Qy 61 GTCAATTCAAGGTGTGAATGTGAAACCAAGTAAAGCTTATACGATGTCGACAGTATGCG 120
Db |||||||
Qy 61 GTCAATTCAAGGTGTGAATGTGAAACCAAGTAAAGCTTATACGATGTCGACAGTATGCG 120
Db |||||||
Qy 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTCTCGAAAA 180
Db |||||||
Qy 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTCTCGAAAA 180
Db |||||||
Qy 181 CCGCGGAAAAGTGGAGCGGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db |||||||
Qy 181 CCGCGGAAAAGTGGAGCGGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db |||||||
Qy 241 AACAACTGGCGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db |||||||
Qy 241 AACAACTGGCGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db |||||||
Qy 301 ACGCGCGTTCGCAAAATTTGTCGCGCGATTTAAATCTCGCGCCGATCAATCGGCGTGCAGCG 360
Db |||||||
Qy 301 ACGCGCGTTCGCAAAATTTGTCGCGCGATTTAAATCTCGCGCCGATCAATCGGCGTGCAGCG 360
Db |||||||
Qy 361 TGGTGGTTCGATGTTAGAACGAGCGCGTTCGAGCCCTGAAAGCGCGGTGCAACAATC 420
Db |||||||
Qy 361 TGGTGGTTCGATGTTAGAACGAGCGCGTTCGAGCCCTGAAAGCGCGGTGCAACAATC 420
Db |||||||
Qy 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATCTCGCTGATGATCAACAGGATGCA 480
Db |||||||
Qy 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATCTCGCTGATGATCAACAGGATGCA 480
Db |||||||
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGCTTATTTCTTGATGCTCTGACGAGA 540
Db |||||||
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGCTTATTTCTTGATGCTCTGACGAGA 540
Db |||||||
Qy 541 CACCCATCAACAGTATTTATTTTCCCAATGAAGACGTTACGGGATGCGGCTGGGAGCATC 600
Db |||||||

D	b	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
Q	y	601	TGGTCGCATTGGGTACCAGCAAATCGCGCTGTTAGCGGCCCATTAAGTTCTGTCTCGG	660
D	b	601	TGGTCGCATTGGGTACCAGCAAATCGCGCTGTTAGCGGCCCATTAAGTTCTGTCTCGG	660
Q	y	661	CGCGTCTGGCTTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
D	b	661	CGCGTCTGGCTTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Q	y	721	CGGAACGGGAAGCGCACCTGGAGTGCATGTCCGTTTTCAACAACCAATGCAATTCGTGA	780
D	b	721	CGGAACGGGAAGCGCACCTGGAGTGCATGTCCGTTTTCAACAACCAATGCAATTCGTGA	780
Q	y	781	ATGAGGGCATCGTCCCACCTGCGATCTGGTTGGCCAACGATCAGATGCGTGGCGGCGAA	840
D	b	781	ATGAGGGCATCGTCCCACCTGCGATCTGGTTGGCCAACGATCAGATGCGTGGCGGCGAA	840
Q	y	841	TGCGGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATACG	900
D	b	841	TGCGGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATACG	900
Q	y	901	ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACACCAATCAAAACGATTTC	960
D	b	901	ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACACCAATCAAAACGATTTC	960
Q	y	961	GCCTGCTGGGGCAAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
D	b	961	GCCTGCTGGGGCAAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Q	y	1021	AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAGAAAACCACTGGCGGCCAATA	1080
D	b	1021	AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAGAAAACCACTGGCGGCCAATA	1080
Q	y	1081	CGCAACCGCCTCTCCCCCGCGGTTGGCGGATTCATTAATCGAGTGGCGACGACAGTTT	1140
D	b	1081	CGCAACCGCCTCTCCCCCGCGGTTGGCGGATTCATTAATCGAGTGGCGACGACAGTTT	1140
Q	y	1141	CCCGACTGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTACCTCACTCATTAG	1200
D	b	1141	CCCGACTGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTACCTCACTCATTAG	1200
Q	y	1201	GCACAAATCTCATGTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGCGG	1260
D	b	1201	GCACAAATCTCATGTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGCGG	1260
Q	y	1261	TCAGCAGCCATCGGAAGCTGTGATATGGCTGTGCGAGTGTGTAATCACTGCGATAATTCG	1320
D	b	1261	TCAGCAGCCATCGGAAGCTGTGATATGGCTGTGCGAGTGTGTAATCACTGCGATAATTCG	1320
Q	y	1321	TGTCGCTCAAGGCGCATCTCCGTTCTCGATAATGTTTTTGGCGCGACATCAATACGGTT	1380
D	b	1321	TGTCGCTCAAGGCGCATCTCCGTTCTCGATAATGTTTTTGGCGCGACATCAATACGGTT	1380
Q	y	1381	CTGGCAATAATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
D	b	1381	CTGGCAATAATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Q	y	1441	ATTGTGAGCGGATHAACAAATTTACAGAGGAAAACGACGTCGGTTAGGTGTTTTACAGA	1500
D	b	1441	ATTGTGAGCGGATHAACAAATTTACAGAGGAAAACGACGTCGGTTAGGTGTTTTACAGA	1500
Q	y	1501	GCATCTCACAAACGAGGACCATGATTATGAAAACTGAAGAGTAACTGGTAACTCTGG	1560
D	b	1501	GCATCTCACAAACGAGGACCATGATTATGAAAACTGAAGAGTAACTGGTAACTCTGG	1560
Q	y	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTTCGGTAAGAAATTCGAGAAAT	1620
D	b	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTTCGGTAAGAAATTCGAGAAAT	1620
Q	y	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCCGATTAACATGGAAGAGAAATTCACAGGTT	1680
D	b	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCCGATTAACATGGAAGAGAAATTCACAGGTT	1680

[illegible]

RESULT 9

```

US-10-343-859-9
; Sequence 9, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; TITLE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9191
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-muts
; OTHER INFORMATION: plasmid
; US-10-343-859-9

```

Query Match	81.8%	Score 2698;	DB 19;	Length 9191;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 2701;	Conservative	0;	Mismatches	5;
			Indels	0;
			Gaps	0;

Qy	1	CCGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
Db	1	CCGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
Qy	61	GTCAATTCCAGGGTGGTGAATGTG	AAACCAAGTAACGTTATACGATCTCGCAGAGTATCGCG	120
Db	61	GTCAATTCCAGGGTGGTGAATGTG	AAACCAAGTAACGTTATACGATCTCGCAGAGTATCGCG	120
Qy	121	GTGTCTCTTATCAGACCGTTTCC	CGCGTGGTGAACCAAGCCAGCCACGTCGTTTCTGCGAAAA	180
Db	121	GTGTCTCTTATCAGACCGTTTCC	CGCGTGGTGAACCAAGCCAGCCACGTCGTTTCTGCGAAAA	180
Qy	181	CGCGGAAAAAGTGGAAACGGCGAT	GGCGAGCTGAATTCACATCCCAACCGCTGGCAC	240
Db	181	CGCGGAAAAAGTGGAAACGGCGAT	GGCGAGCTGAATTCACATCCCAACCGCTGGCAC	240
Qy	241	AACAACTGGCGGGGAAAACAGTCG	TGTTGATTTGGCGTTGCCACCTTCAGTCTGGGCCCTGCG	300
Db	241	AACAACTGGCGGGGAAAACAGTCG	TGTTGATTTGGCGTTGCCACCTTCAGTCTGGGCCCTGCG	300
Qy	301	ACGGCCGTCGCAAAATTCGCGGG	CGAATTAATCTCGCGCCGATCAACTGGGTGGCCAGCG	360
Db	301	ACGGCCGTCGCAAAATTCGCGGG	CGAATTAATCTCGCGCCGATCAACTGGGTGGCCAGCG	360
Qy	361	TGTTGGTGTGATGGTGAAGCGA	AGCGCGTGAAGCCTGTAAAGCGCGGTGCAACAATC	420
Db	361	TGTTGGTGTGATGGTGAAGCGA	AGCGCGTGAAGCCTGTAAAGCGCGGTGCAACAATC	420
Qy	421	TTCTCGCGCAACGCGTCAGTGG	CGTGATCATTAATCTCGCTGATGACCAAGGATGCCA	480
Db	421	TTCTCGCGCAACGCGTCAGTGG	CGTGATCATTAATCTCGCTGATGACCAAGGATGCCA	480
Qy	481	TTGTGTGGAAAGCTGCCTGCAC	TAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCA	540
Db	481	TTGTGTGGAAAGCTGCCTGCAC	TAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCA	540
Qy	541	CACCCATCAACAGTATTATTTTCT	CCCATGAAGACGGTACGCGATCGGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTTCT	CCCATGAAGACGGTACGCGATCGGCGTGGAGCATC	600
Qy	601	TGTTGCGCATTGGGTCAACAGCA	AAATTCGGCTGTTTAGCGGGCCCAATTAAGTTCTGTCTCGG	660

601	TG	TCGATCATGGGCTACCGACAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660
661	Qy	CGCGTCTGCGCTCTGGCTGGCTGGCATAAAATATCTCACTCGCAATCAAATTCAGCCCATAG	720
661	Db	CGCGTCTGCGCTCTGGCTGGCTGGCATAAAATATCTCACTCGCAATCAAATTCAGCCCATAG	720
721	Qy	CGGAA CGGGAAAGCGCATGAGTGGCCATGTC CGGTTTTTCAACAAA CCAATGCAAAATGCTGA	780
721	Db	CGGAA CGGGAAAGCGCATGAGTGGCCATGTC CGGTTTTTCAACAAA CCAATGCAAAATGCTGA	780
781	Qy	ATGAGGCAATCGTTTCCCACTCGCATGCTGGTGCACACGATCAGATGGCGCTGGGCGCAA	840
781	Db	ATGAGGCAATCGTTTCCCACTCGCATGCTGGTGCACACGATCAGATGGCGCTGGGCGCAA	840
841	Qy	TGGCGGCATATACCGAGTCCGCGCTGGCTGGTGGGATATCTCGGTAGTGGGATACG	900
841	Db	TGGCGGCATATACCGAGTCCGCGCTGGCTGGTGGGATATCTCGGTAGTGGGATACG	900
901	Qy	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAA CCAATCAAA CAGGATTTTC	960
901	Db	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAA CCAATCAAA CAGGATTTTC	960
961	Qy	GCCTGCTGGGGCAAA CCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
961	Db	GCCTGCTGGGGCAAA CCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
1021	Qy	AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACACACCTCGCGCCCAATA	1080
1021	Db	AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACACACCTCGCGCCCAATA	1080
1081	Qy	CGAAACCGCTCTCTCCCGCGGTTGGCCGATTAATTAATGCAAGTGGCACGACAGGTTT	1140
1081	Db	CGAAACCGCTCTCTCCCGCGGTTGGCCGATTAATTAATGCAAGTGGCACGACAGGTTT	1140
1141	Qy	CCGCACTGGAAGCGGGCAGTGAAGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG	1200
1141	Db	CCGCACTGGAAGCGGGCAGTGAAGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG	1200
1201	Qy	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCA CGGTGCACCAATGCTTCTGGCG	1260
1201	Db	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCA CGGTGCACCAATGCTTCTGGCG	1260
1261	Qy	TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTCAGGTGTAATCACTGTCATAATTCG	1320
1261	Db	TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTCAGGTGTAATCACTGTCATAATTCG	1320
1321	Qy	TGTCGCTCAAGGGCAGCTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACCGGTT	1380
1321	Db	TGTCGCTCAAGGGCAGCTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACCGGTT	1380
1381	Qy	CTGSCAAATATCTGAAATCAGCTGTTTGACAAATTAATCATCGGTCCGTAAATGTTGGGA	1440
1381	Db	CTGSCAAATATCTGAAATCAGCTGTTTGACAAATTAATCATCGGTCCGTAAATGTTGGGA	1440
1441	Qy	ATTGTGAGCGGATAACAATTTTCA CACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCACGA	1500
1441	Db	ATTGTGAGCGGATAACAATTTTCA CACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCACGA	1500
1501	Qy	GCATTTACCAACAAAGAACATAGATTATGAAAATCTGAAGAGGTAAACTGTTAATCTGG	1560
1501	Db	GCATTTACCAACAAAGAACATAGATTATGAAAATCTGAAGAGGTAAACTGTTAATCTGG	1560
1561	Qy	ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620
1561	Db	ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620
1621	Qy	ACCGGAATTAAGTCA CCGTTGAGCATCCGGATAAATCGGAAGAGAAATTTCCCA CAGGTT	1680
1621	Db	ACCGGAATTAAGTCA CCGTTGAGCATCCGGATAAATCGGAAGAGAAATTTCCCA CAGGTT	1680
1681	Qy	CGGCGAATCGGCGATGGCCCTGACATTTATCTTCTGGGCA CACGACCGCTTTGGTGGCTAC	1740
1681	Db	CGGCGAATCGGCGATGGCCCTGACATTTATCTTCTGGGCA CACGACCGCTTTGGTGGCTAC	1740

```
QY 1741 GCTCAATCTGGCTGTGTGGCTGAATACACCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAATACACCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACTCGGATGCGGTAGCTTTACACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACTCGGATGCGGTAGCTTTACACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
QY 1921 GAGATCCGCGCTCGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCGCGCTCGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCTACTTACCTGGCGCTGATTTGCTGCTGACGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCTACTTACCTGGCGCTGATTTGCTGCTGACGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAAGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAAGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGGCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGGCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACACGCAAGAGTGAATATGCTGTAACGGTACTCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGCAAGAGTGAATATGCTGTAACGGTACTCCGACCTTC 2280
QY 2281 AAGGTCACACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACCGCCGAGT 2340
Db 2281 AAGGTCACACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACCGCCGAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAAGTCTTACGAGGAAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAAGTCTTACGAGGAAAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGAGGATTTTCAAGATTC 2700
QY 2701 GGATCC 2706
Db 2701 GGATCC 2706
```

RESULT 10
US-10-263-153-40
; Sequence 40, Application US/10263153

```
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 7112
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3189)
; OTHER INFORMATION: pMBP-c2X-Toxop30del11 (52-214aa)
; US-10-263-153-40
```

```
Query Match 81.5%; Score 2690.4; DB 18; Length 7112;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGTTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCAATTCAGGTTGTAATGTGAACCCAGTACGTTATACGATGTGCGAGATATGCCG 120
Db 61 GTCAATTCAGGTTGTAATGTGAACCCAGTACGTTATACGATGTGCGAGATATGCCG 120
QY 121 GTGCTCTTTATCAGACCGTTTCCCGGTGTGAACCCAGCGCCAGCCAGCTTCTGCGAAAA 180
Db 121 GTGCTCTTTATCAGACCGTTTCCCGGTGTGAACCCAGCGCCAGCCAGCTTCTGCGAAAA 180
QY 181 CCGCGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTTACATTTCCCAACCGGTGGCAC 240
Db 181 CCGCGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTTACATTTCCCAACCGGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACACAGTCTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACSCGCGTCCGCAAAATTTGTCGGGCGAATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACSCGCGTCCGCAAAATTTGTCGGGCGAATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGTCGATGGTAGAACGCGGTGCGAGCCCTGTAAGCGCGGTGCACAATC 420
Db 361 TGGTGTGTGTCGATGGTAGAACGCGGTGCGAGCCCTGTAAGCGCGGTGCACAATC 420
QY 421 TTCTCCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACGAGATGCCA 480
Db 421 TTCTCCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACGAGATGCCA 480
QY 481 TTGCTGTGAAGCTGCTGCACTAATATTTTCCCGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACTAATATTTTCCCGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGAGTACGCGCTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGAGTACGCGCTGGCGGTGGAGCATC 600
QY 601 TGGTTCGATTTGGGTACACGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTTCGATTTGGGTACACGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CCGCTGTGCTGTGGCTGGCTGGCATTAATATCTCACTCGCAATCAATTTACGCCGATAG 720
```

Db 661 |||||CGGCTCTGGCTGGCTGGAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTGCACACGATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTGCACACGATCAGATGCGCTGGCGGCA 840
Qy 841 TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGTAGTGGGATACG 900
Db 841 TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATCAACACGAGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATCAACACGAGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCACGGCGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCACGGCGTGA 1020
Qy 1021 AGGCAATCAGCTGTGCCCGTCTCACTGGTGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGCCCGTCTCACTGGTGAAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCACGACGAGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCACGACGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTGTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTGTAGTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTCAGAGCTTATCATGACATGACGGTGACCAATGCTTCTGGG 1260
Db 1201 GCACAAATCTCATGTTTCAGAGCTTATCATGACATGACGGTGACCAATGCTTCTGGG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGTAATCACTGCATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGCTTCGGAATATGTTTTTGGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGCTTCGGAATATGTTTTTGGCGGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATCTGAAATGAGCTGTTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTTAGTGTGTTTCAAGA 1500
Db 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTTAGTGTGTTTCAAGA 1500
Qy 1501 GCACCTCAACAAAGGACCATAGATATGAAAACCTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTCAACAAAGGACCATAGATATGAAAACCTGAAGAGGTAAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTTCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTAACAACGCAAGCTGATGTTTACCCGATCGCTTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTAACAACGCAAGCTGATGTTTACCCGATCGCTTT 1860
Qy 1861 GAAGGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGAATAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGAATAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCTGATCTTCACTGCGCGTGAATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCTGATCTTCACTGCGCGTGAATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGAAATGCAAGACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGAAATGCAAGACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGTTCCACATCGACACCGACCAAGTGAATTAATGTTGTAACGGTACTCGGACCTTC 2280
Db 2221 GCATGTTCCACATCGACACCGACCAAGTGAATTAATGTTGTAACGGTACTCGGACCTTC 2280
Qy 2281 AAGGTTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGGTTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGGTTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Qy 2461 TTGGCAAGAGTCCACGATTTTGGCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCAAGAGTCCACGATTTTGGCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCGAGCGTGTGAGTGTGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCGAGCGTGTGAGTGTGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Qy 2641 AACACAAACAATAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAAAATTC 2700
Db 2641 AACACAAACAATAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAAAATTC 2700

RESULT 11
US-10-263-153-30
; Sequence 30, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.

/ TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
/ TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
/ FILE REFERENCE: 6984.US.01
/ CURRENT APPLICATION NUMBER: US/10/263,153
/ CURRENT FILING DATE: 2002-10-22
/ NUMBER OF SEQ ID NOS: 74
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 30
/ LENGTH: 7259
/ TYPE: DNA
/ ORGANISM: Toxoplasma gondii
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1528)...(3336)
/ OTHER INFORMATION: pMBP-c2X-Toxop30del14del18 (83-294aa)
US-10-263-153-30

Query Match 81.5%; Score 2690.4; DB 18; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CGACACCATCGAATGGTCAAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
DB 1 CGACACCATCGAATGGTCAAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
QY 61 CTCGAATTCAGGTGCTGAATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCCG 120
DB 61 CTCGAATTCAGGTGCTGAATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCGCCGAGCCAGCTTTCTCGGAAA 180
DB 121 GTGCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCGCCGAGCCAGCTTTCTCGGAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGAACAGTCTGTGATTTGGGTGGCCACTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGAACAGTCTGTGATTTGGGTGGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTGCACAAATTTGTCGGCGGATTAATCTCGCGCGATCACTGGGTGCCAGCG 360
DB 301 ACGCGCGTGCACAAATTTGTCGGCGGATTAATCTCGCGCGATCACTGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGTGAACGAAGCGCGTGAAGCGCTGTAAGCGCGGTGCACAATC 420
DB 361 TGGTGGTTCGATGTGAACGAAGCGCGTGAAGCGCTGTAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGGTCAAGTGGGCTGATCAATTAATCTCGCGCGATCACTGGGTGCCA 480
DB 421 TTCTCGCGCAACGGTCAAGTGGGCTGATCAATTAATCTCGCGCGATCACTGGGTGCCA 480
QY 481 TTGCTGGGAAGCTGCTGCAATATGTTCCGGGTTATTTCTTGATGTTCTTGACCGAGA 540
DB 481 TTGCTGGGAAGCTGCTGCAATATGTTCCGGGTTATTTCTTGATGTTCTTGACCGAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCAAGAGCGTACGCGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTAATTTCTCCCAAGAGCGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGGCATTTGGGTCAACAGCAATTCGGTGTGGCGGCGCCATTAAGTTCTGCTCGG 660
DB 601 TGGTGGCATTTGGGTCAACAGCAATTCGGTGTGGCGGCGCCATTAAGTTCTGCTCGG 660
QY 661 CGGCTCGGCTTGGGTGGGTGCGATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGGCTCGGCTTGGGTGGGTGCGATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGGCATGTCGGGTTTCAACCAATCAATGCTGA 780
DB 721 CGGAACGGGAAGGCGACTGGAGTGGCATGTCGGGTTTCAACCAATCAATGCTGA 780

QY 781 ATGAGGCGCATGTTCCCACTGCGATGCTGGTGGCCAAAGCATCAGATGGCGCTGGCGCAA 840
DB 781 ATGAGGCGCATGTTCCCACTGCGATGCTGGTGGCCAAAGCATCAGATGGCGCTGGCGCAA 840
QY 841 TGGCGGCCATTAACGAGTCCGGGCTGGCGGTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB 841 TGGCGGCCATTAACGAGTCCGGGCTGGCGGTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCAATCAACAGATTTTC 960
DB 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCAATCAACAGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACGAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACGAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
QY 1021 AGGCAATACGATGTTGCCGCTCTCACTGTGTAAGAAAAAACCACCTCGCGCCCAATA 1080
DB 1021 AGGCAATACGATGTTGCCGCTCTCACTGTGTAAGAAAAAACCACCTCGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGATTCATTAATGCACTGGGACGACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGATTCATTAATGCACTGGGACGACAGGTTT 1140
QY 1141 CCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAAGTGAAGTGAAGTGAAG 1200
DB 1141 CCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAAGTGAAGTGAAGTGAAG 1200
QY 1201 GCACAATTCATGTTTGAAGCTTATCATCGACTGCAAGGTGCAACATGCTTCTGGCG 1260
DB 1201 GCACAATTCATGTTTGAAGCTTATCATCGACTGCAAGGTGCAACATGCTTCTGGCG 1260
QY 1261 TCAGGACGACCATCGGAAGCTGCTGATGGTGGCAGTGGTAAATCACTGCAATATTCG 1320
DB 1261 TCAGGACGACCATCGGAAGCTGCTGATGGTGGCAGTGGTAAATCACTGCAATATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCCTGGATTAATGTTTTCGCCCGACATCAATAAGGTT 1380
DB 1321 TGTGCTCAAGCGCACTCCCGTTCCTGGATTAATGTTTTCGCCCGACATCAATAAGGTT 1380
QY 1381 CTGGCAATATTTCTGAATGAGCTGTGCAATTAATCATCGCTCGTATATGTTGTTGA 1440
DB 1381 CTGGCAATATTTCTGAATGAGCTGTGCAATTAATCATCGCTCGTATATGTTGTTGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTCCGTTTGAAGTGTTCACCA 1500
DB 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTCCGTTTGAAGTGTTCACCA 1500
QY 1501 GCACTTCAACCAAGGACCATAGATTAATGAAACTGAAAGAGGTAAACTGGTAAATCTCG 1560
DB 1501 GCACTTCAACCAAGGACCATAGATTAATGAAACTGAAAGAGGTAAACTGGTAAATCTCG 1560
QY 1561 ATTAAACGGGATAAAGCGTATACCGTCTCGGTGAGTAAAGTAAAGTAAAGTAAAGT 1620
DB 1561 ATTAAACGGGATAAAGCGTATACCGTCTCGGTGAGTAAAGTAAAGTAAAGTAAAGT 1620
QY 1621 ACGGAATTAAGTCAACCGTTCAGCATCCGATTAATGAAACTGAAAGAGGAAATTCACCAAGT 1680
DB 1621 ACGGAATTAAGTCAACCGTTCAGCATCCGATTAATGAAACTGAAAGAGGAAATTCACCAAGT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTCTGCGGACACGACCGCTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTCTGCGGACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGAACAGCGTTCAGGACAGCTGTAT 1800
DB 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGAACAGCGTTCAGGACAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATCGCGTACGTTTAAACGCGCAAGCTGATTGCTTACCAGTCCGCTGT 1860
DB 1801 CGGTTTACCTGGGATCGCGTACGTTTAAACGCGCAAGCTGATTGCTTACCAGTCCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATTAACAAAGATCTGCTGCCGAACCCGCCAAACCTCGGAA 1920

```

1861 GAAGGCTTATCGCTGATTTATTAACAAGATCTGCTGCCGAACCCGCCAAAACCTTGGAA 1920
1921 GAGATCCCGCGCTCGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
1921 GAGATCCCGCGCTCGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTACTTACCTCGCGCGCTGATGCTGCTGACGGGGGTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTACCTCGCGCGCTGATGCTGCTGACGGGGGTATGCGTTCAAG 2040
2041 TATGAAAAACGCAAGTACGACATTAAGACGTGGGCGTGGATAACGCTGGGCGGAAAGCG 2100
2041 TATGAAAAACGCAAGTACGACATTAAGACGTGGGCGTGGATAACGCTGGGCGGAAAGCG 2100
2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAGCGGAAACATGAATGCAGACACCGATTAC 2160
2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAGCGGAAACATGAATGCAGACACCGATTAC 2160
2161 TCCATCGCAGAAGCTGCGCTTTAATAAAGCGGAAACATGAATGCAGACACCGATTAC 2220
2161 TCCATCGCAGAAGCTGCGCTTTAATAAAGCGGAAACATGAATGCAGACACCGATTAC 2220
2221 GCATGCTCAACATCGACACGCAAGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
2221 GCATGCTCAACATCGACACGCAAGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
2281 AAGGTCACACCATCAACACCGTTGCTGGCGTGTGAGCGCAGGTATTAACCGCCCGCAGT 2340
2281 AAGGTCACACCATCAACACCGTTGCTGGCGTGTGAGCGCAGGTATTAACCGCCCGCAGT 2340
2341 CGAACAAGAGCTGGGAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGCTG 2400
2341 CGAACAAGAGCTGGGAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGCTG 2400
2401 GAAGCGGTTAATAAAGACAAACCGCTGGGCGTGGAGCGTGAAGTCTTACGAGGAAG 2460
2401 GAAGCGGTTAATAAAGACAAACCGCTGGGCGTGGAGCGTGAAGTCTTACGAGGAAG 2460
2461 GAAGCGGTTAATAAAGACAAACCGCTGGGCGTGGAGCGTGAAGTCTTACGAGGAAG 2520
2461 GAAGCGGTTAATAAAGACAAACCGCTGGGCGTGGAGCGTGAAGTCTTACGAGGAAG 2520
2521 CGAACAATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGATGATCAACGCC 2580
2521 CGAACAATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGATGATCAACGCC 2580
2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
2641 AACAAACAACAATAACAATAACAACCACTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
2641 AACAAACAACAATAACAATAACAACCACTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
```

RESULT 12

```

US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 35
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528) ... (3399)
; OTHER INFORMATION: pMBP-c2X-Toxop30del10 (52-284aa)
US-10-263-153-35
```

```

Query Match      81.5%; Score 2690.4; DB 18; Length 7322;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGCTGCAAAACCTTTCCGCGTATGCGCATGATAGCCCGCGGAAGAGA 60
Db 1 CCGACACCATCGAATGCTGCAAAACCTTTCCGCGTATGCGCATGATAGCCCGCGGAAGAGA 60

Qy 61 GTCAATTCCAGGCTGTAATGTGAACCCAGTAAAGTATATAGATGTCGACAGTATGCG 120
Db 61 GTCAATTCCAGGCTGTAATGTGAACCCAGTAAAGTATATAGATGTCGACAGTATGCG 120

Qy 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCCAGCGCAGCCAGCTTTCTGCAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCCAGCGCAGCCAGCTTTCTGCAAAA 180

Qy 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGGGCAACAGTCTGCTGATGCGGCTTGCACCTCCAGTCTGCGCCCTGC 300
Db 241 AACAACTGGCGGGGCAACAGTCTGCTGATGCGGCTTGCACCTCCAGTCTGCGCCCTGC 300

Qy 301 ACGCGCGCTCGCAAAATGCTCGCGCGATTAATCTCGCGCGATCAATCGGTGCGCAGCG 360
Db 301 ACGCGCGCTCGCAAAATGCTCGCGCGATTAATCTCGCGCGATCAATCGGTGCGCAGCG 360

Qy 361 TGGTGGTTCGATGATGAACGAGCGCGTGAAGCGCTGTAAGCGCGGTGCAATC 420
Db 361 TGGTGGTTCGATGATGAACGAGCGCGTGAAGCGCTGTAAGCGCGGTGCAATC 420

Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATATCCGCTGATGATCAACGAGTGA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATATCCGCTGATGATCAACGAGTGA 480

Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGCGCTTATTTCTTGATGCTCTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGCGCTTATTTCTTGATGCTCTGACACAGA 540

Qy 541 CACCCATCAACAGTATTAATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

Qy 601 TGGTGGCATTTGGGTACACGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTCTGTCGG 660
Db 601 TGGTGGCATTTGGGTACACGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTCTGTCGG 660

Qy 661 CGCGTCTGCGTCTGCTGCGTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCTGCGTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Qy 721 CGGAACGGGAAGGGGACCTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGGGACCTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780

Qy 781 ATGAGGGCATCTGTCCTCCCATCGCATGCTGGTGTGCAACGATCAGATGCGGCTGGCGCAA 840
Db 781 ATGAGGGCATCTGTCCTCCCATCGCATGCTGGTGTGCAACGATCAGATGCGGCTGGCGCAA 840

Qy 841 TGCGGCCATTAACGAGTCCGGGCTGCGGCTGGTGGCGGATATCTCGTAGTGGGATAGC 900
Db 841 TGCGGCCATTAACGAGTCCGGGCTGCGGCTGGTGGCGGATATCTCGTAGTGGGATAGC 900
```



```
QY 901 ACGATACCGAAGCAGCTCATGTATATATCCCGCGTTAAACCACTCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGCAGCTCATGTATATATCCCGCGTTAAACCACTCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACAGCGTGAGACCGCTTGCTGCAACTCTCTCAGGCCACAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACAGCGTGAGACCGCTTGCTGCAACTCTCTCAGGCCACAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTCAAAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTCAAAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGAATTCATTAATGAGCTGGCAGCACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGAATTCATTAATGAGCTGGCAGCACAGGTTT 1140
QY 1141 CCGGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCACGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCACGGTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCGTGATATGTTTTCGGCGGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCGTGATATGTTTTCGGCGGACATCAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTATGAACTGAAGAGTAACTGATGTTATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTATGAACTGAAGAGTAACTGATGTTATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAACTGGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAACTGGAAGAGAAATTCGCCACAGTT 1680
QY 1681 CGCGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGCAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGGCAAGCTGATGTTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGGCAAGCTGATGTTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGACGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGACGCGCTGATGTTCAAC 1980
```

```
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGANTTCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGANTTCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTAGTACGACATTAAAGACGCTGGCGGTGGATAACGCTGGCGCAAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTAGTACGACATTAAAGACGCTGGCGGTGGATAACGCTGGCGCAAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAACAAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAACAAACACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCTCTTAAATAAGGCGAAACAGCGATGACCATCAACGGCGCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTCTTAAATAAGGCGAAACAGCGATGACCATCAACGGCGCCGTGG 2220
QY 2221 GCATGGTCCAACATCGACACACGCAAAAGTGAATTTATGTTGTAACGGTACTGCCACCTTC 2280
Db 2221 GCATGGTCCAACATCGACACACGCAAAAGTGAATTTATGTTGTAACGGTACTGCCACCTTC 2280
QY 2281 AAGGGTCAACATCGACACACGCTTTCGTTGGCGTGTGAGCGCAGGTATTAACGGCCAGT 2340
Db 2281 AAGGGTCAACATCGACACACGCTTTCGTTGGCGTGTGAGCGCAGGTATTAACGGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAGGTCTG 2400
QY 2401 GAAGCGGTTTAATAAAGACAAACCGCTGGTGGCTAGCGCTGAACTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTTAATAAAGACAAACCGCTGGTGGCTAGCGCTGAACTTTACGAGGAAGAG 2460
QY 2461 TTGCGGAAACATCCAGTATTGCGGCACCATCGAACCAACCGCTGAGCTGAACTGAAATCATG 2520
Db 2461 TTGCGGAAACATCCAGTATTGCGGCACCATCGAACCAACCGCTGAGCTGAACTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
QY 2581 GCAGCGGTGCTGAGCTGTCGATGAAAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCAGCGGTGCTGAGCTGTCGATGAAAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATTAACATAACAAATTAACAAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATTAACATAACAAATTAACAAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
```

RESULT 13

```
US-10-263-153-25
; Sequence 25, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 7352
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3429)
```


OTHER INFORMATION: pMBP-c2X-Toxop30del4C (52-294aa)									
US-10-263-153-25									
Query Match 81.5%; Score 2690.4; DB 18; Length 7352;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;									
QY	1	CGGACACCATCGAATGGTGGCAAAACCTTTTCGGGATATGGCATGATAGCGCCCGGAAGAGA	60						
DB	1	CGGACACCATCGAATGGTGGCAAAACCTTTTCGGGATATGGCATGATAGCGCCCGGAAGAGA	60						
QY	61	GTCAATTACAGGTGGTGAATGTGAACACCAAGTATACGATGTCGACAGTATGCGG	120						
DB	61	GTCAATTACAGGTGGTGAATGTGAACACCAAGTATACGATGTCGACAGTATGCGG	120						
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGTTCCTGCGAAAA	180						
DB	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGTTCCTGCGAAAA	180						
QY	181	CGCGGAAAGTGAAGCGGGGATGGCGGAGCTGAATTACATTTCCCAACCGCTGGCAC	240						
DB	181	CGCGGAAAGTGAAGCGGGGATGGCGGAGCTGAATTACATTTCCCAACCGCTGGCAC	240						
QY	241	AACRACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
DB	241	AACRACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
QY	301	ACGCGCGTTCGCAAAATTTGCGGGGATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG	360						
DB	301	ACGCGCGTTCGCAAAATTTGCGGGGATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG	360						
QY	361	TGGTGGTTCGATGTAGAACGAAAGCGCGTGAAGCCCTGTAAGCGCGGTGCACAAATC	420						
DB	361	TGGTGGTTCGATGTAGAACGAAAGCGCGTGAAGCCCTGTAAGCGCGGTGCACAAATC	420						
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGACCAAGGATGCCA	480						
DB	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGACCAAGGATGCCA	480						
QY	481	TTGCTGTGGAAGCTGCCTGCATTAATGTTTCGGCGGTATTTCTGATGTCTCTGACCAGA	540						
DB	481	TTGCTGTGGAAGCTGCCTGCATTAATGTTTCGGCGGTATTTCTGATGTCTCTGACCAGA	540						
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCATGCGGCGTGGAGCATC	600						
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCATGCGGCGTGGAGCATC	600						
QY	601	TGGTGGCATTTGGTCCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG	660						
DB	601	TGGTGGCATTTGGTCCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG	660						
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
QY	721	CGGAACGGGAAGCGACGTGAGTGCATGTCCGTTTTTCAACAAACCATGCAAAATGCTGA	780						
DB	721	CGGAACGGGAAGCGACGTGAGTGCATGTCCGTTTTTCAACAAACCATGCAAAATGCTGA	780						
QY	781	ATGAGGGCATGTTTCCCACTGGATGCTGGTTGGCCACGATCAGATGGCGCTGGGGCNA	840						
DB	781	ATGAGGGCATGTTTCCCACTGGATGCTGGTTGGCCACGATCAGATGGCGCTGGGGCNA	840						
QY	841	TGCGGCGCATATACCGAGTCCGGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATAGC	900						
DB	841	TGCGGCGCATATACCGAGTCCGGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATAGC	900						
QY	901	ACGATACCGGAAGACAGCTCATGTATATCCCGCGGTAAACCAATCAACAGGATTTTC	960						
DB	901	ACGATACCGGAAGACAGCTCATGTATATCCCGCGGTAAACCAATCAACAGGATTTTC	960						
QY	961	GCCTGCTGGGGCAACACGAGTGGACCGGTTGCTGCAACTCTCAGGGCCAGGCGGTGA	1020						
DB	961	GCCTGCTGGGGCAACACGAGTGGACCGGTTGCTGCAACTCTCAGGGCCAGGCGGTGA	1020						

DB	961	GCCTGCTGGGGCAACACGAGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020						
QY	1021	AGGCAATCAGCTGTGTCCCGTCTCACTGTGTGAAGAAAAACCACTTGGCGCCCAATA	1080						
DB	1021	AGGCAATCAGCTGTGTCCCGTCTCACTGTGTGAAGAAAAACCACTTGGCGCCCAATA	1080						
QY	1081	CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCAGCTGCACACAGGTTT	1140						
DB	1081	CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCAGCTGCACACAGGTTT	1140						
QY	1141	CCGACTGGAAGAGGGGAGTGCAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTAG	1200						
DB	1141	CCGACTGGAAGAGGGGAGTGCAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTAG	1200						
QY	1201	GCACAATTTCTATGTTTGACAGCTTATCATCGATCGACGGTGCAACCAATGCTTCTGGG	1260						
DB	1201	GCACAATTTCTATGTTTGACAGCTTATCATCGATCGACGGTGCAACCAATGCTTCTGGG	1260						
QY	1261	TCAGCAGCCATCGGAGCTGTGTATGCGTGTGAGGTGCTGAGTCACTGCATTAATTCG	1320						
DB	1261	TCAGCAGCCATCGGAGCTGTGTATGCGTGTGAGGTGCTGAGTCACTGCATTAATTCG	1320						
QY	1321	TGTGCTCAAGCGGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380						
DB	1321	TGTGCTCAAGCGGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380						
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATATGTTGGA	1440						
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATATGTTGGA	1440						
QY	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500						
DB	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500						
QY	1501	GCATTTCACCAACAGGACCATAGATTAATGAAAACTGAAAGAGGTAAACTGTTAATCTGG	1560						
DB	1501	GCATTTCACCAACAGGACCATAGATTAATGAAAACTGAAAGAGGTAAACTGTTAATCTGG	1560						
QY	1561	ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620						
DB	1561	ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620						
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCCAGGTT	1680						
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCCAGGTT	1680						
QY	1681	CGGCAACTGCGCATGCGCTTGACATTAATCTTCTGGGCACACGACCGCTTGTGGTGCTAC	1740						
DB	1681	CGGCAACTGCGCATGCGCTTGACATTAATCTTCTGGGCACACGACCGCTTGTGGTGCTAC	1740						
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGACAAAGCTGAT	1800						
DB	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGACAAAGCTGAT	1800						
QY	1801	CGTTTACCTGGATGCGGTAGTTTAAACGGCAAGCTGATGTTTACCGGATCCCTGTT	1860						
DB	1801	CGTTTACCTGGATGCGGTAGTTTAAACGGCAAGCTGATGTTTACCGGATCCCTGTT	1860						
QY	1861	GAAGGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920						
DB	1861	GAAGGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920						
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980						
DB	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980						
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATGTTGCTGACGCGGGTTATGCGTTCAAG	2040						
DB	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATGTTGCTGACGCGGGTTATGCGTTCAAG	2040						
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACCGTGGCGTGGATTAACGCTGGCGGAAAGCG	2100						
DB	2041	TATGAAAAACGCAAGTACGACATTAAGACCGTGGCGTGGATTAACGCTGGCGGAAAGCG	2100						

```
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTCGTGTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAACTGCTTTAATAAGGCGAAACAGCGATACCAATCAACGCGCGGTGG 2220
DB 2161 TCCATCGCAGAACTGCTTTAATAAGGCGAAACAGCGATACCAATCAACGCGCGGTGG 2220
QY 2221 GCATGGTCCAAACATCCACACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
DB 2221 GCATGGTCCAAACATCCACACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCGTTGGCTGCTGAGCGCAGGTATTAAACGCGCCAGT 2340
DB 2281 AAGGGTCAACCATCCAAACCGTTCGTTGGCTGCTGAGCGCAGGTATTAAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGCGAAAGAGTTCCTCGAAACTATCTGTAAGTGAAGGTCG 2400
DB 2341 CCGAACAAAGAGCTGCGAAAGAGTTCCTCGAAACTATCTGTAAGTGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
DB 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGTTATTCGCGCACCATGGAAGAAACGCGCAAGGTAATCATG 2520
DB 2461 TTGGCGAAGATCCACGTTATTCGCGCACCATGGAAGAAACGCGCAAGGTAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGTTGATGCGGTGCGTACTGCGGTGATCAAGCC 2580
DB 2521 CCGAACATCCCGCAGATGTCGCTTCTGTTGATGCGGTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
DB 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACA 2700
DB 2641 AACAAACAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACA 2700
```

RESULT 14

```
US-10-263-153-20
; Sequence 20, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxo30del3C (52-300aa)
US-10-263-153-20
```

Query Match

```
Best Local Similarity 81.5%; Score 2690.4; DB 18; Length 7370;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 CCGACACCATCGAATGGTGTGCAAAAACCTTTTCGCGGTATGGCATGTAGCGCCCGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGTGCAAAAACCTTTTCGCGGTATGGCATGTAGCGCCCGAAGAGA 60
QY 61 GTCAATTCAGAGGTGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATGCCG 120
DB 61 GTCAATTCAGAGGTGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCCTTATCAGACCGTTCGCGCGTGTGTAACACGAGCCAGCCAGTTCCTCGGAAAA 180
DB 121 GTGTCCTTATCAGACCGTTCGCGCGTGTGTAACACGAGCCAGCCAGTTCCTCGGAAAA 180
QY 181 CGCGGGAAGTGGAGCGCGGATGCGGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
DB 181 CGCGGGAAGTGGAGCGCGGATGCGGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAACAGTCTGTTGCTGATTGGCGTTCACCTCCAGTCTGCCCTGC 300
DB 241 AACAACTGGCGGGCAACAGTCTGTTGCTGATTGGCGTTCACCTCCAGTCTGCCCTGC 300
QY 301 ACGCCGCTGCGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGCCGCTGCGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTTAGAACGAGCGCGTGCAGGCTGTAAGCGCGCGGTGCAATC 420
DB 361 TGGTGGTGTGATGTTAGAACGAGCGCGTGCAGGCTGTAAGCGCGCGGTGCAATC 420
QY 421 TTCTCGCGCAACCGTCAAGTGGCTGATCAATTAATCTCGCTGGATGACACGAGTGCCA 480
DB 421 TTCTCGCGCAACCGTCAAGTGGCTGATCAATTAATCTCGCTGGATGACACGAGTGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAAATTTTCGCGCGTATTTCTTGATGTCTCGACAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACTAAATTTTCGCGCGTATTTCTTGATGTCTCGACAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTAGCGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTAGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGGATTTGGTCAACGAGCAATCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTGGATTTGGTCAACGAGCAATCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGGCTCTGCGTCTGGCTGCTGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGGCTCTGCGTCTGGCTGCTGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGCATCGAGTGGCCATGTCGCGTTCCTCAACAAACCATGCAAACTCGA 780
DB 721 CGGAACGGGAAGCGCATCGAGTGGCCATGTCGCGTTCCTCAACAAACCATGCAAACTCGA 780
QY 781 ATGAGGCGATCGTCCCACTGCGATGCTGGTGGCCAAAGATGAGTGGCGTGGGCGCAA 840
DB 781 ATGAGGCGATCGTCCCACTGCGATGCTGGTGGCCAAAGATGAGTGGCGTGGGCGCAA 840
QY 841 TGGCGGCCATTACCGAGTCCGCGCTGCGGTTGGTGGCGATATCTCGGTAGTGGATAG 900
DB 841 TGGCGGCCATTACCGAGTCCGCGCTGCGGTTGGTGGCGATATCTCGGTAGTGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCTGAGCCGCTGCTGCAACTCTCTCAGGGCCAGCGCGTGA 1020
DB 961 GCCTGCTGGGCAACACAGCTGAGCCGCTGCTGCAACTCTCTCAGGGCCAGCGCGTGA 1020
QY 1021 AGGGCAATCAGTGTTCGCGCTCTCACTGTGTAAGAAAGAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGTGTTCGCGCTCTCACTGTGTAAGAAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCAATTAATGAGCTGCGACGACAGGTTT 1140
```

```
Db 1081 |||||CGAAAACGCGCTCTCCCGCGGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATTTCTATGTTTGACAGCTTATCATGACTGCGACGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTTCTATGTTTGACAGCTTATCATGACTGCGACGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTGTAATCACTGCATAAATTCG 1320
Qy 1321 TGTGCGCTCAAGCGGCACCTCCGTTCTCGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCGCTCAAGCGGCACCTCCGTTCTCGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTGTTGCACAAATTAATCATCGGCTCGTATATGTGTGGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGCACAAATTAATCATCGGCTCGTATATGTGTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTACACAGAAACAGCCAGTCGTTTAGTGTTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTACACAGAAACAGCCAGTCGTTTAGTGTTTTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAAGGTAAACTGCTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAAGGTAAACTGCTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATAAATCGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATAAATCGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 GCGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCCGTAGTTTACACGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCCGTAGTTTACACGCAAGCTGATGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGAAGCGTGGCGGTGGATAACCGTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGAAGCGTGGCGGTGGATAACCGTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACCAACATGAATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACCAACATGAATGACAGACCGGATTAC 2160
Qy 2161 TCCATCGCAAGAGCTGCCTTTAATTAAGCGGAAACAGGATGACCATCAACGGCCCGGTGG 2220
Db |||||
```

```
Db 2161 TCCATCGCAGAGCTGCCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTAATGGTTAAACGGTACTCCGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTAATGGTTAAACGGTACTCCGACCTTC 2280
Qy 2281 AAGGTCNAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCNAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG 2400
Qy 2401 GAAGCGGTTAATAAAGCAAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTACGAGAAAG 2460
Db 2401 GAAGCGGTTAATAAAGCAAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTACGAGAAAG 2460
Qy 2461 TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAACCGCCAGAAAGGTTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAACCGCCAGAAAGGTTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGGTCGTACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCGTACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACACACACATTAACAATAACAACCAACCTCCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACACACACACATTAACAATAACAACCAACCTCCGGATCGAGGAGGATTTTCAGAAATTC 2700
```

RESULT 15

```
US-10-263-153-61
; Sequence 61, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984 US 01
; CURRENT APPLICATION NUMBER: US/10/263.153
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-ToxoP30MIX1
US-10-263-153-61
```

```
Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCCCGGGAAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCCCGGGAAGA 60
Qy 61 GTCAATTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGTTATCCGACAGTATCGCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGTTATCCGACAGTATCGCG 120
```

QY	121	GTGTCTCTTATCAGACCGTTTCCGGCTGTGTGAACGAGGCCAGCCAGCTTCTCTGCGAAA	180
Db	121	GTGTCTCTTATCAGACCGTTTCCGGCTGTGTGAACGAGGCCAGCCAGCTTCTGCGAAA	180
QY	181	CGCGGGAAGTGGAAAGCGCGGATGGCGAGCTGAATACATTCACACCGCGTGGCAC	240
Db	181	CGCGGGAAGTGGAAAGCGCGGATGGCGAGCTGAATACATTCACACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTGCAGAAATGTTCGGCGGATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCGTGCAGAAATGTTCGGCGGATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG	360
QY	361	TGCTGTGTGCGATGGTAGAACGAGCGCGTGAAGCCTGTGAAGCGCGGTGCACAATC	420
Db	361	TGCTGTGTGCGATGGTAGAACGAGCGCGTGAAGCCTGTGAAGCGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGTGC	480
Db	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGTGC	480
QY	481	TTGCTGTGAAGCTGCTGCACTAATGTTCGGCGTTATTTCTTGATGTCTCTGACCCAGA	540
Db	481	TTGCTGTGAAGCTGCTGCACTAATGTTCGGCGTTATTTCTTGATGTCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATC	600
QY	601	TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCG	660
Db	601	TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCG	660
QY	661	CGCGTCTGGCTGGCTGGGCAATAATCTCACTCGCAATCAAAATCAGCCGATAG	720
Db	661	CGCGTCTGGCTGGCTGGGCAATAATCTCACTCGCAATCAAAATCAGCCGATAG	720
QY	721	CGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCGATCGTTCCCACTCGCATGCTGTGTCACAGATCAGATGCGTGGCGCAA	840
Db	781	ATGAGGCGATCGTTCCCACTCGCATGCTGTGTCACAGATCAGATGCGTGGCGCAA	840
QY	841	TGCGCGCATTTACGAGTCCGGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACGAGTCCGGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCAACCATCAACAGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCAACCATCAACAGATTTTC	960
QY	961	GCCTGCTGGGCAACAGCGTGAACCGCTTGTGCAACTCTCTCAGGCGCAAGCGGTGA	1020
Db	961	GCCTGCTGGGCAACAGCGTGAACCGCTTGTGCAACTCTCTCAGGCGCAAGCGGTGA	1020
QY	1021	AGGCAATACAGCTGTTCGCGCTCTCACTGTGAAAGAAAAAACCCCTCGCGCCCAATA	1080
Db	1021	AGGCAATACAGCTGTTCGCGCTCTCACTGTGAAAGAAAAAACCCCTCGCGCCCAATA	1080
QY	1081	CGCAACCGGCTCTCCCGCGGTTGGCCGATTCATTAATGCAAGTGGCAAGCAGGTTT	1140
Db	1081	CGCAACCGGCTCTCCCGCGGTTGGCCGATTCATTAATGCAAGTGGCAAGCAGGTTT	1140
QY	1141	CCGACTGGAAAGCGGCGAGTGAAGCAAGCAATTAATGCTGAGTTAGTCACTCATTTAG	1200
Db	1141	CCGACTGGAAAGCGGCGAGTGAAGCAAGCAATTAATGCTGAGTTAGTCACTCATTTAG	1200

QY	1201	GCAAAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGCTGCACCAATCTTCTGGCG	1260
Db	1201	GCAAAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGCTGCACCAATCTTCTGGCG	1260
QY	1261	TCAGCAGCCCATCGGAAGCTGTGTATGGCTGTGTCAGGTCGTAATCATCTGCATAAATTCG	1320
Db	1261	TCAGCAGCCCATCGGAAGCTGTGTATGGCTGTGTCAGGTCGTAATCATCTGCATAAATTCG	1320
QY	1321	TGTCGCTCAAGCGCGACCTCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCATACCGTT	1380
Db	1321	TGTCGCTCAAGCGCGACCTCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCATACCGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTGTGTGAACAATTAATCATCGGCTCGTATAATGTGTGA	1440
Db	1381	CTGCAAAATTTCTGAAATGAGCTGTGTGAACAATTAATCATCGGCTCGTATAATGTGTGA	1440
QY	1441	ATTCTGAGCGGATAAACAATTTACACAGGAACAGCCAGTCCGTTTAAAGTGTGTTTACGA	1500
Db	1441	ATTCTGAGCGGATAAACAATTTACACAGGAACAGCCAGTCCGTTTAAAGTGTGTTTACGA	1500
QY	1501	GCATTCACCAACAGGACCATAGATTTATGAAAATCTGAAGAGTAAATCTGGTAATCTGG	1560
Db	1501	GCATTCACCAACAGGACCATAGATTTATGAAAATCTGAAGAGTAAATCTGGTAATCTGG	1560
QY	1561	ATTAACCGCGATAAAGCTATTAACGGTCTCGCTGAAGTCTGGTAAGAAATTCGAGAAGAT	1620
Db	1561	ATTAACCGCGATAAAGCTATTAACGGTCTCGCTGAAGTCTGGTAAGAAATTCGAGAAGAT	1620
QY	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACCAAGTT	1680
Db	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACCAAGTT	1680
QY	1681	GGCGCACTGGCGATGGCCCTGACATTTCTCTGGGCAACGACCGCTTTGGTGGCTAC	1740
Db	1681	GGCGCACTGGCGATGGCCCTGACATTTCTCTGGGCAACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCGGACAAAGCGTTCCAGAGCAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCGGACAAAGCGTTCCAGAGCAAGCTGTAT	1800
QY	1801	CGTTTACTGGGATGCGGTGATTAACAACGCAAGCTGATTTACCCGATCGCTGTT	1860
Db	1801	CGTTTACTGGGATGCGGTGATTAACAACGCAAGCTGATTTACCCGATCGCTGTT	1860
QY	1861	GAGCGTATCGCTGATTTTAAACAAGATCTGCTGCCGAACCCGCAAAAACCTGGGAA	1920
Db	1861	GAGCGTATCGCTGATTTTAAACAAGATCTGCTGCCGAACCCGCAAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGGTGATTTGCTGTGACCGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGGCGGTGATTTGCTGTGACCGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAACGGCAAGTACGACATTAAGACCGTGGCGTGAATACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAAACGGCAAGTACGACATTAAGACCGTGGCGTGAATACGCTGGCGCGAAAGCG	2100
QY	2101	GCTCTGACCTTCTCGGTTGACCTGATTAATAACCAACATGATGACACCGATTTAC	2160
Db	2101	GCTCTGACCTTCTCGGTTGACCTGATTAATAACCAACATGATGACACCGATTTAC	2160
QY	2161	TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
QY	2221	GCATGGTCCAAATCGACACCAAGTGAATTTATGCTGTAACGCTACTGCCAGCTTC	2280
Db	2221	GCATGGTCCAAATCGACACCAAGTGAATTTATGCTGTAACGCTACTGCCAGCTTC	2280
QY	2281	AAGGCTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT	2340

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-18
Perfect score: 3300
Sequence: 1 ccgcacacatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_ges1:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	23.6	794	5	BQ751936
2	770.6	23.4	779	7	CK394523
3	768	23.3	769	7	CK118014
4	762	23.1	793	5	BQ751655
5	755.4	22.9	806	5	BQ751220
6	684.8	20.8	689	6	CB863541
7	676.4	20.5	712	1	AL038548
8	676	20.5	710	8	AY080106
9	647.4	19.6	660	1	AL044483
10	641	19.4	641	4	BJ063872
11	615.8	18.7	624	6	CB863814
12	612	18.5	794	1	AL045353
13	603	18.3	752	9	CL655575
14	589.8	17.9	608	8	AY080094
15	586.4	17.8	589	1	AL037742
16	581.2	17.6	501	1	AV594538
17	579	17.5	579	1	AL043868
18	570	17.3	782	7	CK394246
19	569	17.2	774	1	AL045337
20	563.2	17.1	585	9	CL658017
21	562.6	17.0	723	1	AL0379416
22	549.4	16.6	753	1	AL045341
23	546.8	16.6	782	1	AL037051
24	545.4	16.5	547	6	CA895956

C 25	545	16.5	648	7	CK394349	CK394349	hggada2D0
C 26	540.2	16.4	569	1	AL039077	AL039077	DKF2p566G
C 27	534.4	16.2	644	6	CB862524	CB862524	HH04B08u
C 28	534	16.2	734	1	AL048427	AL048427	DKF2p588H
C 29	529	16.0	740	1	AL042909	AL042909	DKF2p434J
C 30	528	16.0	720	6	CA881974	CA881974	K0994C10-
C 31	523.2	15.9	528	1	AL038811	AL038811	DKF2p566O
C 32	519.4	15.7	521	6	CA890032	CA890032	B0156F03-
C 33	519	15.7	607	6	CB862099	CB862099	HH06A02Y
C 34	516.4	15.6	518	6	CA886082	CA886082	B0124F04-
C 35	510.2	15.5	686	1	AL044407	AL044407	DKF2p434D
C 36	509.4	15.4	511	6	CA895436	CA895436	B0192A07-
C 37	498	15.1	498	1	AL039076	AL039076	DKF2p566G
C 38	494.8	15.0	1025	1	AL038025	AL038025	DKF2p566K
C 39	487.6	14.8	498	7	CK394397	CK394397	hggada2H1
C 40	477.6	14.5	491	6	CB862715	CB862715	HH03F14u
C 41	465.8	14.1	679	1	AL039128	AL039128	DKF2p566K
C 42	462.8	14.0	490	1	AL039589	AL039589	DKF2p434D
C 43	458	13.9	458	6	CA887583	CA887583	B0137H04-
C 44	456.4	13.8	471	1	AL039649	AL039649	DKF2p434G
C 45	449.2	13.6	612	6	CB863409	CB863409	HH04K05w

ALIGNMENTS

RESULT 1
BQ751936
LOCUS BQ751936 794 bp mRNA linear EST 18-JUL-2002
DEFINITION EST632499 DSCT Colletotrichum trifolii cDNA clone pDSCT8-67, mRNA
sequence.
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1. (bases 1 to 794)
Samad, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
Cheung, F. and Fraser, C.M.
ESTs from mycelia of Colletotrichum trifolii race 1
Other ESTs: EST632498
Unpublished (2002)
Contact: Deborah A. Samad
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debby@puccini.crl.umn.edu
TIGR sequence name: MTSAH67TV More information is available at:
www.medicago.org
Seq primer: (gca Ata Cga Ctc Act Ata 999 C).
Location/Qualifiers
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSCT8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after
inoculation) grown in liquid culture (cutin minimal medium
containing 2%glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"

FEATURES
source

1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSCT8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after
inoculation) grown in liquid culture (cutin minimal medium
containing 2%glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"

/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
EcoRI; isolate: 2ap2 ; cDNA was prepared from polyA+
enriched RNA the cDNA was ligated into Lambda gIII from
Stratagene and packaged using Gigapack packaging extracts.
An aliquot of the amplified library was used to transduce.

E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match	23.6%;	Score 778;	DB 5;	Length 794;	
Best Local Similarity	98.7%;	Pred. No. 3.4e-219;			
Matches 784;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;	
QY	3	GACACATCGATGTGC	AAACCTTTCCG	TGATGCGCATGATAGCGCCGGAAGAGT	62
Db	1	GACATTTTGAATGCG	CAAAACCTTTCCG	TGATGCGCATGATAGCGCCGGAAGAGT	60
QY	63	CAATTTCAGGTTGGT	GAATGTGAAC	CAGTAAAGTATACGATGTCGACAGATGTCGGGT	122
Db	61	CAATTTCAGGTTGGT	GAATGTGAAC	CAGTAAAGTATACGATGTCGACAGATGTCGGGT	120
QY	123	GTCTCTTATCAGAC	CGTTTCCGCGTGT	GAACACGAGCCAGCGTTTCTGCGAAACG	182
Db	121	GTCTCTTATCAGAC	CGTTTCCGCGTGT	GAACACGAGCCAGCGTTTCTGCGAAACG	180
QY	183	CGGGAAGAGTGGAG	CGCGATGCGG	AGCTGAATTAACATTCACACCGCGTGACAA	242
Db	181	CGGGAAGAGTGGAG	CGCGATGCGG	AGCTGAATTAACATTCACACCGCGTGACAA	240
QY	243	CAACTGGCGGCAAC	AGTCTGTTGCTG	ATGCGCTGTCACCTCCAGTCTGCGCTGCAC	302
Db	241	CAACTGGCGGCAAC	AGTCTGTTGCTG	ATGCGCTGTCACCTCCAGTCTGCGCTGCAC	300
QY	303	CGCGCTGCGCAAA	TGTCGCGCGAT	TAAATCTCGCGCGATCAACTGGGTGCCAGCGTG	362
Db	301	CGCGCTGCGCAAA	TGTCGCGCGAT	TAAATCTCGCGCGATCAACTGGGTGCCAGCGTG	360
QY	363	GTGGTGTGATGTA	GAACGAGCGGT	CGAAGCTGTAAGCGCGGTGTCACATCTT	422
Db	361	GTGGTGTGATGTA	GAACGAGCGGT	CGAAGCTGTAAGCGCGGTGTCACATCTT	420
QY	423	CTCGCGAACCGT	CAGTGGCTGAT	CATTAACATTCGCTGGATGACCGAGTGCAT	482
Db	421	CTCGCGAACCGT	CAGTGGCTGAT	CATTAACATTCGCTGGATGACCGAGTGCAT	480
QY	483	GCTGTGNAAGTGC	TGACATAATGTT	CCGCGTTATTTCTTGATGTCCTGACCAACA	542
Db	481	GCTGTGNAAGTGC	TGACATAATGTT	CCGCGTTATTTCTTGATGTCCTGACCAACA	540
QY	543	CCCATCAACAGT	ATTATTTCTCC	ATGAACGCTACGCACTGGGGTGGACATCTG	602
Db	541	CCCATCAACAGT	ATTATTTCTCC	ATGAACGCTACGCACTGGGGTGGAGCATCTG	600
QY	603	GTGCAATTGGTTC	ACCAAGCAATCG	CGCTGTTAGCGGCGCCATTAAGTTCCTCGCGG	662
Db	601	GTGCAATTGGTTC	ACCAAGCAATCG	CGCTGTTAGCGGCGCCATTAAGTTCCTCGCGG	660
QY	663	CGTCTGGTCTGG	CTGGTGCATAATA	TCTCACTCGCAATCAAAATTCAGCCGATAGCG	722
Db	661	CGTCTGGTCTGG	CTGGTGCATAATA	TCTCACTCGCAATCAAAATTCAGCCGATAGCG	720
QY	723	GAACGGGAAGCG	AGCTGGAGTGC	ATGTCGCGTTTCAACAAACCATGCAATGCTGAAT	782
Db	721	GAACGGGAAGCG	AGCTGGAGTGC	ATGTCGCGTTTCAACAAACCATGCAATGCTGAAT	780
QY	783	GAGGGCATCGTTC	796		
Db	781	GAGGGCATCGTTC	794		

RESULT 2
CK394523/c
LOCUS
DEFINITION hggada4D08 Gland Cell Amplified cDNA Library Heterodera glycines 779 bp mRNA linear EST 30-DEC-2003

cdna, mRNA sequence.
CK394523
CK394523.1 GI:40389794
EST.
Heterodera glycines
Heterodera glycines
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchida;
Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
1 (bases 1 to 779)
Wang,X., Allen,R., Gao,B., Goellner,M., Maier,T., Baum,T.,
Hussey,R. and Davis,E.
Submission: Wang, Allen, Gao, Goellner, Maier, Baum, Hussey, Davis
Unpublished (2003)
Contact: Tom Maier
Department of Plant Pathology, Baum Lab
Iowa State University
351 Bessey Hall, Ames, IA 50011, USA
Tel: 515-294-8854
Fax: 515-294-9420
Email: tmaier@iastate.edu
Heterodera glycines Gland Cell Amplified cDNA Library, single pass
sequence.
Location/Qualifiers
1..779
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell Amplified cDNA Library"
/notes="Organ: gland cell; Vector: pSportII"

ORIGIN

Query Match	23.4%;	Score 770.6;	DB 7;	Length 779;	
Best Local Similarity	99.4%;	Pred. No. 5.3e-217;			
Matches 773;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;	
QY	87	CCAGTAACGTTATAC	GATGTCGAGAGT	ATGCGCGTGTCTCTTATCAGACCGTTTCCCGC	146
Db	779	CCAGTAACGTTATAC	GATGTCGAGAGT	ATGCGCGTGTCTCTTATCAGACCGTTTCCCGC	720
QY	147	GTGGTGAACAGCG	CAGCCACGTTTCT	CGAAAAACGCGGAAAAAGTGGAAAGCGGCGATG	206
Db	719	GTGGTGAACAGCG	CAGCCACGTTTCT	CGAAAAACGCGGAAAAAGTGGAAAGCGGCGATG	660
QY	207	GCGGAGCTGAATT	TACATTCCTCAAC	CGCGTGGCACAACTCTGCGGGGCAACAGTCGTTG	266
Db	659	GCGGAGCTGAATT	TACATTCCTCAAC	CGCGTGGCACAACTCTGCGGGGCAACAGTCGTTG	600
QY	267	CTGATTGGCGGTT	GGCCACCTCCAGT	CTGGCCCTGCACGCGCGTCCGAAATTTGTCGGCGG	326
Db	599	CTGATTGGCGGTT	GGCCACCTCCAGT	CTGGCCCTGCACGCGCGTCCGAAATTTGTCGGCGG	540
QY	327	ATTTAAATCTCG	CGCCCATCACTGG	GTGTCAGGTGTCGATGTGAGAACGAAAGC	386
Db	539	ATTTAAATCTCG	CGCCCATCACTGG	GTGTCAGGTGTCGATGTGAGAACGAAAGC	480
QY	387	GCGGTGGAAGC	CTGTAAGCGCGT	GCAATCTTCTCGCGCAACGCGTCAGTGGGCTG	446
Db	479	GCGGTGGAAGC	CTGTAAGCGCGT	GCAATCTTCTCGCGCAACGCGTCAGTGGGCTG	420
QY	447	ATCATTAATCAT	CCGCTGGATGAC	CCAGGATGCCATTTGCTGTGGAAGCTGCTGCATTAAT	506
Db	419	ATCATTAATCAT	CCGCTGGATGAC	CCAGGATGCCATTTGCTGTGGAAGCTGCTGCATTAAT	360
QY	507	GTTCGGCGGTT	ATTCTTCTGATGT	CTCTGACGACACCCATCAACAGATATATTTCTCC	566
Db	359	GTTCGGCGGTT	ATTCTTCTGATGT	CTCTGACGACACCCATCAACAGATATATTTCTCC	300
QY	567	CATGAAGACGGT	ACGCGACTGGG	CGGTGGAGCATCTGGTCGATTGGGTCAACGAGAAATC	626
Db	299	CATGAAGACGGT	ACGCGACTGGG	CGGTGGAGCATCTGGTCGATTGGGTCAACGAGAAATC	240

COMMENT

Other ESTs: EST632217
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAF59TV More information is available at:
 www.medicago.org
 Seq primer: (gtA AtA CgA Ctc Act AtA 99g C).

FEATURES

source

1. .793
 Location/Qualifiers

/organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST6-59"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DST"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match
 Best Local Similarity 23.1%; Score 762; DB 5; Length 793;
 Matches 777; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
 3 GACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCGGAAGAGT 62
 1 GACATCTTCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCGGAAGAGT 60
 63 CAATTCAGGGTGGTGAATGTGAACACCGTAAAGTTATACGATGCGCAGAGTATGCCGGT 122
 61 CAATTCAGGGTGGTGAATGTGAACACCGTAAAGTTATACGATGCGCAGAGTATGCCGGT 120
 123 GTCTCTTATCAGACCGTTTCCGCGGTGTGAACACCGTAAAGTTATACGATGCGCAGAGTATGCCGGT 182
 121 GTCTCTTATCAGACCGTTTCCGCGGTGTGAACACCGTAAAGTTATACGATGCGCAGAGTATGCCGGT 180
 183 CGGGAAAAAGTGGAAAGCGCGGATCGCGAGTGAATTAATTCACATTCCTCAACCGCGTGGACAA 242
 181 CGGGAAAAAGTGGAAAGCGCGGATCGCGAGTGAATTAATTCACATTCCTCAACCGCGTGGACAA 240
 243 CAATCGCGGCAACAGTGTGCTGATGGCGGTGCGACCTCCAGTTCGGCCCTGCGAC 302
 241 CAATCGCGGCAACAGTGTGCTGATGGCGGTGCGACCTCCAGTTCGGCCCTGCGAC 300
 303 GCGCGGTGCAAAATGTTCGCGGGGATTAATTCGCGCGGATCAATCACTGGGTGCGACCGT 362
 301 GCGCGGTGCAAAATGTTCGCGGGGATTAATTCGCGCGGATCAATCACTGGGTGCGACCGT 360
 363 GTGGTGTGATGTGAGAACGAGCGGTGAGAGCTGTAAGCGCGGTGAGCAATCTT 422
 361 GTGGTGTGATGTGAGAACGAGCGGTGAGAGCTGTAAGCGCGGTGAGCAATCTT 420
 423 CTGCGCAACCGGTGAGTGGGCTGATCAATTAATTCCTCGGTGATGATCACCAGGATGCCATT 482
 421 CTGCGCAACCGGTGAGTGGGCTGATCAATTAATTCCTCGGTGATGATCACCAGGATGCCATT 480

QY 483 GCTGTGGAAGCTGCCTGCACATAATGTTCCGCGGTTATTTCTTGATGTTCTCTGACGACA 542
 DB 481 GCTGTGGAAGCTGCCTGCACATAATGTTCCGCGGTTATTTCTTGATGTTCTCTGACGACA 540
 QY 543 CCCATCACAGTATATTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATCTG 602
 DB 541 CCCATCACAGTATATTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATCTG 600
 QY 603 GTCGATTGGTTCACACGAAATCGCGCTGTAGCGGCCCATTAAGTTCTGTCTCGCG 662
 DB 601 GTCGATTGGTTCACACGAAATCGCGCTGTAGCGGCCCATTAAGTTCTGTCTCGCG 660
 QY 663 GCTCTGGCTCTGGCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
 DB 661 GCTCTGGCTCTGGCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720
 QY 723 GAACGGGAAGCGCACTGGAGTGCATGTCGGTTCAGTTCACAAACCATGCAAAATGCTGAAT 782
 DB 721 GAACGGGAAGCGCACTGGAGTGCATGTCGGTTCAGTTCAGTTCGCGTTTCA--CAAACATGCAAAATGCTGAAT 778
 QY 783 GAGG 786
 DB 779 GAGG 782

RESULT 5

LOCUS

DEFINITION

BO751220 806 bp mRNA linear EST 18-JUL-2002
 EST631783 DSCT Colletotrichum trifolii cDNA clone pDST3-64, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BO751220 806 bp mRNA linear EST 18-JUL-2002
 EST631783 DSCT Colletotrichum trifolii cDNA clone pDST3-64, mRNA sequence.
 BO751220
 BO751220.1 GI:21906625
 EST
 Colletotrichum trifolii
 Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
 mitosporic Phyllachoraceae; Colletotrichum.
 1 (bases 1 to 806)
 Samac, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
 Cheung, F. and Fraser, C.M.
 ESTs from mycelia of Colletotrichum trifolii race 1
 Unpublished (2002)
 Other ESTs: EST631782

Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org
 Seq primer: (gtA AtA CgA Ctc Act AtA 99g C).

FEATURES

source

1. .806
 Location/Qualifiers

/organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST3-64"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DST"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid

lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK⁺. Aliquots of the ligation were used to transform *E. coli* DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN	Query Match	22.9%	Score 755.4;	DB 5;	Length 806;
	Best Local Similarity	97.4%	Pred. No. 1.8e-212;		
	Matches 790;	Conservative 0;	Mismatches 16;	Indels 5;	Gaps 2;
Qy	3	GACACCATCGAATGGTGC AAAACCTTTCGCGGTATGGCATGATAGCGGCCGGAAGAGAGT	62		
Db	1	GACACTTTCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGGCCGGAAGAGAGT	60		
Qy	63	CAATTCCAGGGTGGTGAATGTGAAACACGATTAAGTTTATACGATGTCGCAAGAGTATGCCCGT	122		
Db	61	CAATTCCAGGGTGGTGAATGTGAAACACGATTAAGTTTATACGATGTCGCAAGAGTATGCCCGT	120		
Qy	123	GTCTCTTTATCAGACCGTTTCCGCGGTGTGAAACACGAGCCAGCCAGCTTTTCTGCGAAAACG	182		
Db	121	GTCTCTTTATCAGACCGTTTCCGCGGTGTGAAACACGAGCCAGCCAGCTTTTCTGCGAAAACG	180		
Qy	183	CGGGAAAAAGTGGAAAGCGGGCGATGCGCGAGCTGAATTACATTTCCCAAACGGGTGGCACAA	242		
Db	181	CGGGAAAAAGTGGAAAGCGGGCGATGCGCGAGCTGAATTACATTTCCCAAACGGGTGGCACAA	240		
Qy	243	CAACTCGCGGGCAACAGTCTGCTGATTTGGCGTTGGCCACTTCAGTCTGGCCCTGCGAC	302		
Db	241	CAACTCGCGGGCAACAGTCTGCTGATTTGGCGTTGGCCACTTCAGTCTGGCCCTGCGAC	300		
Qy	303	CGCGCGTTCGCAAAATGTTCGCGCGGATTAATACTTCGCGCGGATCAACTGGGTGCCAGCGTG	362		
Db	301	CGCGCGTTCGCAAAATGTTCGCGCGGATTAATACTTCGCGCGGATCAACTGGGTGCCAGCGTG	360		
Qy	363	GTGTGTTCGATGGTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCACAATCTT	422		
Db	361	GTGTGTTCGATGGTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCACAATCTT	420		
Qy	423	CTCGCGCAACCGTTCAGTGGCGTGTGATTAACATATCGCTGGATGACACGATGCCATT	482		
Db	421	CTCGCGCAACCGTTCAGTGGCGTGTGATTAACATATCGCTGGATGACACGATGCCATT	480		
Qy	483	GCTGTGGAAGTGCCTGCACATAAGTTCGCGGGTTATTTCTTGATGTCTCTACACAGACA	542		
Db	481	GCTGTGGAAGTGCCTGCACATAAGTTCGCGGGTTATTTCTTGATGTCTCTACACAGACA	540		
Qy	543	CCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCACTGGCGGTGGAGCATCTG	602		
Db	541	CCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCACTGGCGGTGGAGCATCTG	600		
Qy	603	GTCCGATTGGGTACACGCAAAATCGCGCTGTGTAGCGGGGCCCATTAAGTTCTGTCTCGGGG	662		
Db	601	GTCCGATTGGGTACACGCAAAATCGCGCTGTGTAGCGGGGCCCATTAAGTTCTGTCTCGGGG	660		
Qy	663	CGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG	722		
Db	661	CGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG	720		
Qy	723	GAAACGGGAAGCGACTGGAGTGCATATGTCGCGTTTTCACAAACCAATGCTGAAT	782		
Db	721	GAAACGGGAAGCGACTGGAGTGCATATGTCGCGTTTTCACAAACCAATGCTGAAT	780		
Qy	783	GAGGCGATCGTTCCCACTGGCATGCTGGTTG	813		
Db	778	GAGG--CATCGTTCCACTGGCATGCTGGTTG	806		

RESULT 6				
CB863541/c				
LOCUS	CB863541	689 bp	mRNA	linear EST 22-APR-2003
DEFINITION	HH04A08y HH	Hordeum vulgare	cdna clone HH04A08 3-PRIME,	mRNA
				sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1965, Vol. 68, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1978, Vol. 1, No. 2, pp. 1-10
3. The Importance of Parental Involvement	Journal of Educational Psychology	1985, Vol. 77, No. 3, pp. 1-10
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1990, Vol. 93, No. 1, pp. 1-10
5. The Role of the School in the Community	Journal of Educational Research	1995, Vol. 98, No. 1, pp. 1-10
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	2000, Vol. 103, No. 1, pp. 1-10
7. The Role of the Teacher in the 21st Century	Journal of Educational Research	2005, Vol. 108, No. 1, pp. 1-10
8. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	2010, Vol. 113, No. 1, pp. 1-10
9. The Role of the Teacher in the 21st Century	Journal of Educational Research	2015, Vol. 118, No. 1, pp. 1-10
10. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	2020, Vol. 123, No. 1, pp. 1-10

FEATURES

CB863541
CB863541.1 GI:30058100
EST.
Hordeum vulgare
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 689)
Varshey, R.K., Zhang, H., Burton, R., Stein, N., Langridge, P. and
Granger, A.

Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department GenBank
Institute of Plant Genetics and Crop Plant Research (IPK)
Correnstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 689 Std Error: 0.00
Plate: 4 row: 8
Seg primer: SP6.

```

Location/Qualifiers
1. 689
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555198"
/db_xref="taxon:4513"
/clone="HH04A08"
/tissue_type="coleoptile"
/dev_stages="coleoptile, 1 day old"
/lab_host="DH10B"
/clone_lib="HH"
/note="Vector: pSPORT; Site 1: Sal
Site 2: NotI (3-end of cDNA); Due
used blue/white selection for rec
reliable. Average insert size is 1

```

ORIGIN

Query Match	20.8%	Score 684.8;	DB 6;	Length 689;
Best Local Similarity	99.6%	Pred. No. 1.7e-191;		
Matches 686:	Conservative	0: Mismatches	3: Indels	0: Gaps

Qy	21	AAAACTTTCCGGTATGGCATATAGAGCCCGGAAGAGAGTCAATTCAGGGTGGTGAAT	80
Db	689	AAAACTTTCCGGTATGGCATATAGAGCCCGGAAGAGAGTCAATTCAGGTGNGTGAAT	630
Qy	81	GTGAAACCAAGTAAAGCTTATACGATGTGCACAGTATCCGGTGTCCTCTTATCAGACCGTT	140
Db	629	GTGAAACCAAGTAAAGCTTATACGATGTGCACAGTATGCCGGTGTCTCTTATCAGACCGTT	570
Qy	141	TCCCGCGTGTGAACCAAGCCAGCCAGCTTTCTCGGAAAAACGGGAAAAAGTGGAAAGCG	200
Db	569	TCCCGCGTGTGAACCAAGCCAGCCAGCTTTCTCGGAAAAACGGGAAAAAGTGGAAAGCG	510
Qy	201	CGATGGCGGAGCTGAATTACATTCCTCAAACCGCTGGGCAACAACACTGGGGGCAACACAG	260
Db	509	CGATGGCGGAGCTGAATTACATTCCTCAAACCGCTGGGCAACAACACTGGGGGCAACACAG	450
Qy	261	TCGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGCCCTGCACGCGCGCTCGCAAAATTTGC	320
Db	449	TCGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGCCCTGCACGCGCGCTCGCAAAATTTGC	390
Qy	321	CGCGCGATTAAATCTTCGGCGCGATCAACTGGGTGGCAGCGTGGTGGTTCGATGGTATAGAA	380
Db	389	CGCGCGATTAAATCTTCGGCGCGATCAACTGGGTGGCAGCGTGGTGGTTCGATGGTATAGAA	330
Qy	381	CGAAGCGCGCTCGAAGCCTGTAAAGCGGCGGTGCACAAATCTTCTCCGCCAAACCGCGTCAGT	440
Db	329	CGAAGCGCGCTCGAAGCCTGTAAAGCGGCGGTGCACAAATCTTCTCCGCCAAACCGCGTCAGT	270

RESULT 6
CB863541/
LOCUS
DEFINITION

		/mol_type="genomic DNA" /db_xref="taxon:186617" /clone="SI051p36L" /clone_lib="Scripps Pier (La Jolla, CA) uncultured virus community" /note="Marine viruses were isolated from 200 liters of surface seawater using a combination of differential filtration and density-dependent gradient centrifugation. Linker-amplified shotgun libraries were constructed by randomly shearing the total marine viral community DNA, end-repairing, ligating dsDNA linkers to the ends, and amplifying the fragments using Vent DNA polymerase. The resulting fragments were ligated into the pSMART vector and electroporated into MC12 cells (Lucigen; Middleton, WI)."	
ORIGIN		Query Match 20.5%; Score 676; DB 8; Length 710; Best Local Similarity 98.9%; Pred. No. 6.9e-189; Matches 690; Conservative 0; Mismatches 7; Indels 1; Gaps 1;	
Qy	400	GTAAAGCGCGGTGCACAAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAACATATC	459
Db	8	GTGCTGCAGCGGTGCACAAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAACATATC	67
Qy	460	CGCTGGATGACAGATGCCATTGCTGTGGAAGCTGCTGCACATTAATGTTCCGGCGTTAT	519
Db	68	CGCTGGATGACAGATGCCATTGCTGTGGAAGCTGCTGCACATTAATGTTCCGGCGTTAT	127
Qy	520	TTCCTTGATGTCCTGACAGACACCCATCAACAGTATTATTTCTCCCATGAAGACGCTA	579
Db	128	TTCCTTGATGTCCTGACAGACACCCATCAACAGTATTATTTCTCCCATGAAGACGCTA	187
Qy	580	CGGACTGGGCGTGAGAGCATCTGTGCGATTGGGTCAACAGCAATCGCGTGTGTAGCGG	639
Db	188	CGGACTGGGCGTGAGAGCATCTGTGCGATTGGGTCAACAGCAATCGCGTGTGTAGCGG	247
Qy	640	GCCCATTAAGTTCTGCTCTGGGCGCTGTCGCTGTGGCTGGCTGGCATAAATATCTCACTC	699
Db	248	GCCCATTAAGTTCTGCTCTGGGCGCTGTCGCTGTGGCTGGCTGGCATAAATATCTCACTC	307
Qy	700	GCAATCAAAATTCAGCCGATAGCGAAACGGGAGGCGACTGGAGTGCCATGTCGGGTTTC	759
Db	308	GCAATCAAAATTCAGCCGATAGCGAAACGGGAGGCGACTGGAGTGCCATGTCGGGTTTC	367
Qy	760	AACAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTGCCATGCTGTGGTTCGCAACG	819
Db	368	AACAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTGCCATGCTGTGGTTCGCAACG	427
Qy	820	ATCAGATGGCGTGGCGCAATGCGCGCATTAACGAGTCCGGGCTGCGCGTTGGTGGCG	879
Db	428	ATCAGATGGCGTGGCGCAATGCGCGCATTAACGAGTCCGGGCTGCGCGTTGGTGGCG	487
Qy	880	ATATCTCGGTAGTGGATACGACGATACCGAGACAGCTCATGTTATATCCCGCGTTAA	939
Db	488	ATATCTCGGTAGTGGATACGACGATACCGAGACAGCTCATGTTATATCCCGCGTTAA	547
Qy	940	CCACCATCAACAGAGATTTTTCCTGCTGGGCGAAACAGCGTGACCGCTTGTGCAAC	999
Db	548	CCACCATCAACAGAGATTTTTCCTGCTGGGCGAAACAGCGTGACCGCTTGTGCAAC	607
Qy	1000	TCTCTCAGGGCGAGCGGTGAAGGCAATCAAGCTGTGGCCGCTCTCACTGTGGAAGAA	1059
Db	608	TCTCTCAGGGCGAGCGGTGAAGGCAATCANCTGTGGCCGCTCTCACTGTGGAAGAA-AA	666
Qy	1060	AAACCAACCTCGGCCCAATACGAAACCGCTCTCTCC	1097
Db	667	AAACCAACCTCGGCCCAATACGAAACCGNCTTTCCC	704
RESULT 9		AL044483/c	
LOCUS		AL044483 660 bp mRNA linear EST 04-SEP-2003	

DEFINITION	DKFZp434i1102.s1.434 (synonym: htes3) Homo sapiens cDNA clone		
ACCESSION	DKFZp434i1102.3', mRNA sequence.		
VERSION	AL044483		
KEYWORDS	EST.		
SOURCE	AL044483.1 GI:5432701		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 660) Ansorge, W., Benes, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.		
TITLE	EST (Ansorge, Benes, et al.)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: MIPS MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany This is the 3' sequence of the clone insert Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. No r1 sequence available. This clone (DKFZp434i1102) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de. Location/Qualifiers 1..660 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DKFZp434i1102" /tissue_type="testis" /dev_stage="adult" /lab_host="DH10B" /clone_lib="434 (synonym: htes3)" /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"		
FEATURES	source		
ORIGIN	Query Match 19.6%; Score 647.4; DB 1; Length 660; Best Local Similarity 99.7%; Pred No. 2.2e-180; Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;		
QY	85	AACAGTAACGTTATACGATGTCGACAGATATGCGGTGTCCTTTATCAGACCGTTCCC	144
DB	660	AACAGTAACGTTATACGATGTCGACAGATATGCGGTGTCCTTTATCAGACCGTTCCC	601
QY	145	GGTGGTGAACAGCCAGCCAGCTTTCTGCGAAACCGGGGAAAAGTGGNAGCGCGCA	204
DB	600	GGTGGTGAACAGCCAGCCAGCTTTCTGCGAAACCGGGGAAAAGTGGNAGCGCGCA	541
QY	205	TGGCGGAGCTGAATTACATTTCCCAACCGGTGGCAACAACCTGCGCGGCAACAGTGGT	264
DB	540	TGGCGGAGCTGAATTACATTTCCCAACCGGTGGCAACAACCTGCGCGGCAACAGTGGT	481
QY	265	TGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGACCGCGCTCGCAAAATTTGCGCGG	324
DB	480	TGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGACCGCGCTCGCAAAATTTGCGCGG	421
QY	325	CGATTAATCTCGCGCCGATCAACTGGGTGCGAGGTGGTGTGTCGATGTTAGAACGAA	384
DB	420	CGATTAATCTCGCGCCGATCAACTGGGTGCGAGGTGGTGTGTCGATGTTAGAACGAA	361
QY	385	GGGCGGTGCAAGCGCTGTAAGCGGGGTGCACAAATCTTCTCGCGCAACCGCTCAGTGGGC	444
DB	360	GGGCGGTGCAAGCGCTGTAAGCGGGGTGCACAAATCTTCTCGCGCAACCGCTCAGTGGGC	301
QY	445	TGATCATTAATATTCGCTGGATGACACAGGATGCCATTTGCTGTGGAAGCTGCTGCACCTA	504
DB	300	TGATCATTAATATTCGCTGGATGACACAGGATGCCATTTGCTGTGGAAGCTGCTGCACCTA	241
QY	505	ATGTTCCGGCGTTATTTCTTGATGTCTCTGACGACACCCATCAACAGTATTATTTTCT	564

```

Db      240  ATGTTCCGGCGTTATTCTTGATGTTCTCTGACCAGACCCATCAACGATTAATTTCT 181
QY      565  CCATGAAGACGGTAGCGGCTGGGCGTGGAGCATCTGTCGATGGGTACACGACAA 624
Db      180  CCATGAAGACGGTAGCGGCTGGGCGTGGAGCATCTGTCGATGGGTACACGACAA 121
QY      625  TCGCGCTGTTAGCGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGCG 684
Db      120  TCGCGCTGTTAGCGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGCG 62
QY      685  ATAAATATCTCACTCGCAATCAAAATTCACCGCATAGCGGAAACGGGAGCGCATCGAGTG 744
Db      61  ATAAATATCTCACTCGCAATCAAAATTCACCGCATAGCGGAAACGGGAGCGCATCGAGTG 744
QY      745  C 745
Db      1  C 1

RESULT 10
BJ063872/c
LOCUS
DEFINITION
Xenopus laevis (African clawed frog)
ACCESSION
BJ063872
VERSION
BJ063872.1
KEYWORDS
EST.
SOURCE
Xenopus laevis
ORGANISM
Xenopus laevis
REFERENCE
Kitayama, A., Terashima, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tehin@genes.nig.ac.jp
The information of this clone is available through the following URL.
http://xenopus.nibb.ac.jp.
FEATURES
Location/Qualifiers
1. .641
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL077k02"
/tissue_type="whole embryo"
/dev stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"
ORIGIN
Query Match 19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred. NO. 1.7e-178;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      30  CGCGGTATGCATGATACCGCCGAGAGAGATCAATTCAGGTGGTGAATGGAACCA 89
Db      641  CGCGGTATGCATGATACCGCCGAGAGAGATCAATTCAGGTGGTGAATGGAACCA 582
QY      90  GTAAACGTTATACGATGTCGAGAGTATGCGCGGTGTCCTTATCAGACCGTTTCCCGCGTG 149
Db      581  GTAAACGTTATACGATGTCGAGAGTATGCGCGGTGTCCTTATCAGACCGTTTCCCGCGTG 522
QY      150  GTGAACACCGCCAGCCAGCTTTCTCGGAAAAACCGGGGAAAAAGTGAAGCGCGATGGCG 209

```

```

Db      521  GTGAACACCGCCAGCCAGCTTTCTCGGAAAAACCGGGGAAAAAGTGAAGCGCGATGGCG 462
QY      210  GAGCTGAATTACATTCCCAACCGCGTGGCAACAACCTGGCGGGCAACAGTTCGTTGCTG 269
Db      461  GAGCTGAATTACATTCCCAACCGCGTGGCAACAACCTGGCGGGCAACAGTTCGTTGCTG 402
QY      270  ATTGGCGTTGCACTCCAGTCTGGCCCTGCAACGCGCGCTCGCAAAATTGTCGCGCGGATT 329
Db      401  ATTGGCGTTGCACTCCAGTCTGGCCCTGCAACGCGCGCTCGCAAAATTGTCGCGCGGATT 342
QY      330  AAATCTCGCGCGGATCAACTGGGTGCGAGCGTGGTGTGTCGATGGTAGAACGAGCGCG 389
Db      341  AAATCTCGCGCGGATCAACTGGGTGCGAGCGTGGTGTGTCGATGGTAGAACGAGCGCG 282
QY      390  GTCGAAGCCTGTAAAGCGCGGCTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 449
Db      281  GTCGAAGCCTGTAAAGCGCGGCTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 222
QY      450  ATTAACCTATCGGCTGGATGACCAAGATGCCATTTGCTGTGGAAGCTGCTGCACCTAATGTT 509
Db      221  ATTAACCTATCGGCTGGATGACCAAGATGCCATTTGCTGTGGAAGCTGCTGCACCTAATGTT 162
QY      510  CCGCGGTTATTTCTTGATGTTCTGACCAACACCATCAACAGTATTTTCTCCCAT 569
Db      161  CCGCGGTTATTTCTTGATGTTCTGACCAACACCATCAACAGTATTTTCTCCCAT 102
QY      570  GAAGACGGTAGCGGCTGGGCGTGGAGCATCTCGTCGCAATTTGGTCAACAGCAAAATCGCG 629
Db      101  GAAGACGGTAGCGGCTGGGCGTGGAGCATCTCGTCGCAATTTGGTCAACAGCAAAATCGCG 42
QY      630  CTGTTAGCGGGCCCATTAAGTTCTGTCGCGCGCTGCGG 670
Db      41  CTGTTAGCGGGCCCATTAAGTTCTGTCGCGCGCTGTCG 1

RESULT 11
CB863814/c
LOCUS
DEFINITION
HH07B20y HH Hordeum vulgare cdna clone HH07B20 3-PRIME, mRNA
sequence.
ACCESSION
CB863814
VERSION
CB863814.1
KEYWORDS
EST.
SOURCE
Hordeum vulgare
ORGANISM
Hordeum vulgare
REFERENCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 624)
Varshney, R.K., Zhang, H., Burton, R., Stein, N., Langridge, P. and Graner, A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 7 row: B column: 20
Seq primer: SP6.
Location/Qualifiers
1. .624
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="GABI:555471"
/db_xref="taxon:4513"
/clone="HH07B20"
/tissue_type="coleoptile"

```


/dev_stage="coleoptile, 1 day old"

/lab_host="DH10B"

/clone_lib="HH"

/note="Vector: pSPORT; Site_1: SalI (5'-end of cDNA); Site_2: NotI (3'-end of cDNA); Due to the cloning system used blue/white selection for recombinants is not 100 % reliable. Average insert size is 1.3 kb."

ORIGIN

Query Match 18.7%; Score 615.8; DB 6; Length 624;
Best Local Similarity 99.7%; Pred. No. 5.3e-171;
Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GACACATCGATGTCGAAACCTTCGCGTATGGCATGTAGCGCCGGAAGAGT 62
Db |||||||
QY 619 GACACATCGATGTCGAAACCTTCGCGTATGGCATGTAGCGCCGGAAGAGT 560
Db |||||||
QY 63 CAATTTCAGGGTGGTGAATGTAACACAGTAACAGTTATACGATGTCGAGATATCGCGT 122
Db |||||||
QY 559 CAATTTCAGGGTGGTGAATGTAACACAGTAACAGTTATACGATGTCGAGATATCGCGT 500
Db |||||||
QY 123 GTCTCTTATCAGACCGTTTCGCGGTGTAACACAGGCGCAGCAGTTTCTCGGAAAACG 182
Db |||||||
QY 499 GTCTCTTATCAGACCGTTTCGCGGTGTAACACAGGCGCAGCAGTTTCTCGGAAAACG 440
Db |||||||
QY 183 CGGGAAAAGTCGAAGCGCGATGCGGAGCTGAATTAATTCATCCCAACCGCGTGGCACA 242
Db |||||||
QY 439 CGGGAAAAGTCGAAGCGCGATGCGGAGCTGAATTAATTCATCCCAACCGCGTGGCACA 380
Db |||||||
QY 243 CAACCTGGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCGCGCCCTGCAC 302
Db |||||||
QY 379 CAACCTGGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCGCGCCCTGCAC 320
Db |||||||
QY 303 GCGCGTGCAGAAATGTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 362
Db |||||||
QY 319 GCGCGTGCAGAAATGTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 260
Db |||||||
QY 363 GTGGTGTGATGATGTAAGACGCGCTGCGAAGCTGTAAAGCGCGGTGCACAAATCTT 422
Db |||||||
QY 259 GTGGTGTGATGATGTAAGACGCGCTGCGAAGCTGTAAAGCGCGGTGCACAAATCTT 200
Db |||||||
QY 423 CTCGCGCAACGCGTCAGTGGCGTGATTAATCTCGCGCGATCAACTGGGTGCCAGATGCCATT 482
Db |||||||
QY 199 CTCGCGCAACGCGTCAGTGGCGTGATTAATCTCGCGCGATCAACTGGGTGCCAGATGCCATT 140
Db |||||||
QY 483 GCTGTGGAAGCTGCGTGCATTAATGTCGCGCGTATTTCTTGTGATGTCCTCTGACGAGACA 542
Db |||||||
QY 139 GCTGTGGAAGCTGCGTGCATTAATGTCGCGCGTATTTCTTGTGATGTCCTCTGACGAGACA 80
Db |||||||
QY 543 CCATCAACAGATTAATTTCTCCATGAAGACGCTACGCGATCGCGGTGGAGCATCTG 602
Db |||||||
QY 79 CCATCAACAGATTAATTTCTCCATGAAGACGCTACGCGATCGCGGTGGAGCATCTG 20
Db |||||||
QY 603 GTCGCAATGGGTCCAGC 621
Db |||||||
QY 19 GTCGCAATGGGTCCAGC 1

RESULT 12
LOCUS AL045353/c
DEFINITION DKF2p434B075.e1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045353
VERSION AL045353.1 GI:5433506
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 794)
AUTHORS Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

EST (Duesterhoeft, et al.)

Unpublished (1999)

Contact: MIPS

FEATURES

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); sequenced by Qiagen within the cDNA
sequencing consortium of the German Genome Project. r1 sequence
also available.
This clone is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1..794
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p434B075"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 18.5%; Score 612; DB 1; Length 794;
Best Local Similarity 99.7%; Pred. No. 7.6e-170;
Matches 634; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 569 TGAAGACGGTACGGACCTGGGGTGGAGCATCTGTCGCATTTGGTCCACCAGCAATCGC 628
Db |||||||
QY 794 TGAAGACGGTACGGACCTGGGGTGGAG-ATCTGTCGATTTGGTCCACCAGCAATCGC 736
Db |||||||
QY 629 GCTGTTAGCGGCCCATTTAAGTTCTGCTCGCGCGCTCTCGCTCTGGCTGGCTGGCATAA 688
Db |||||||
QY 735 GCTGTTAGCGGCCCATTTAAGTTCTGCTCGCGCGCTCTG-GTCTGGCTGGCTGGCATAA 677
Db |||||||
QY 699 ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGAAACGGGAAGCGACTCGAGTGCCAT 748
Db |||||||
QY 676 ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGAAACGGGAAGCGACTCGAGTGCCAT 617
Db |||||||
QY 749 GTCCGGTTTCAACAAACCATGCAATCTGATGAGGGCATCGTTCCCACTGCCATGCT 808
Db |||||||
QY 616 GTCCGGTTTCAACAAACCATGCAATCTGATGAGGGCATCGTTCCCACTGCCATGCT 557
Db |||||||
QY 809 GGTTCGCCAAGCATCAGATGGCGCTGGCGCAATCGCGCATTTACCGAGTCCGGGCTGGC 868
Db |||||||
QY 556 GGTTCGCCAAGCATCAGATGGCGCTGGCGCAATCGCGCATTTACCGAGTCCGGGCTGGC 497
Db |||||||
QY 869 CGTTGGTCCGATATCTCGGTAGTGGGATAGCAGATACCGAAGACAGCTCATGTTATAT 928
Db |||||||
QY 496 CGTTGGTCCGATATCTCGGTAGTGGGATAGCAGATACCGAAGACAGCTCATGTTATAT 437
Db |||||||
QY 929 CCCGCCGTTAACCCACCATCAAAACAGGATTTTCGCTGCTGGGGCAAAACAGCGTGGACCG 988
Db |||||||
QY 436 CCCGCCGTTAACCCACCATCAAAACAGGATTTTCGCTGCTGGGGCAAAACAGCGTGGACCG 377
Db |||||||
QY 989 CTTGCTGCAACTCTCTCAGGGCCAGCGGTGAAGGGCAATCAGCTGTTGCCGCTCTCACT 1048
Db |||||||
QY 376 CTTGCTGCAACTCTCTCAGGGCCAGCGGTGAAGGGCAATCAGCTGTTGCCGCTCTCACT 317
Db |||||||
QY 1049 GGTGAAAAAGAAAAACACCTCGGCCCAATAACGCAAAACCGCTCTCCCGCGCGGTGGC 1108
Db |||||||
QY 316 GGTGAAAAAGAAAAACACCTCGGCCCAATAACGCAAAACCGCTCTCCCGCGCGGTGGC 257
Db |||||||
QY 1109 CGATTTCATTATGAGCTGGGACAGAGTTTCCCGACTTGGAAAGCGGCGATGAGCGCA 1168
Db |||||||
QY 256 CGATTTCATTATGAGCTGGGACAGAGTTTCCCGACTTGGAAAGCGGCGATGAGCGCA 197
Db |||||||
QY 1169 AGCAATTAATGAGTTAGCTCACTCACTATTAGGCAC 1204
Db |||||||
QY 196 AGCAATTAATGAGTTAGCTCACTCACTATTAGGCAC 161

RESULT 13
CL655575/c
LOCUS
DEFINITION
CL655575
752 bp DNA linear GSS 09-JUL-2004
PRI0124a_C01 - PRI0124a.B21 (752) Mixed stage foetid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL655575
CL655575.1 GI:50134897
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 752)
AppADB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: foetid ends.

FEATURES
Location/Qualifiers
1..752
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage foetid library of P. pacificus
var. California"
/note="Vector: pBpifos-5 Foetid vector"

ORIGIN
Query Match 18.3%; Score 603; DB 9; Length 752;
Best Local Similarity 98.9%; Pred. No. 3.6e-167;
Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 1996 TTCACCTGGCGCTGATTGCTGTGAC-GGGGGTTATCGCTCAAGTATGAAACGGCAA 2054
DB 752 TTCACCTGGCGCTGATTGCTGTGACGGGGTTATCGCTCAAGTATGAAACGGCAA 693
QY 2055 GTACGACATTAAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCGGTCTGACCTTCCT 2114
DB 692 GTACGACATTAAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCGGTCTGACCTTCCT 633
QY 2115 GGTGACCTGATTAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAACG 2174
DB 632 GGTGACCTGATTAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAACG 573
QY 2175 TGCCTTTAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGGCGATGGTCCAAACAT 2234
DB 572 TGCCTTTAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGGCGATGGTCCAAACAT 513
QY 2235 CGACACAGCAAAAGTGAATTATGGTGTAAACGGTACTGCCGACCTTCAAGGGTCAACCATC 2294
DB 512 CGACACAGCAAAAGTGAATTATGGTGTAAACGGTACTGCCGACCTTCAAGGGTCAACCATC 453
QY 2295 CAACCCCTTGTGGCGTGTGAGCGAGGTATTAACGCCGCCAGTCCGACAAAGAGCT 2354
DB 452 CAACCCCTTGTGGCGTGTGAGCGAGGTATTAACGCCGCCAGTCCGACAAAGAGCT 393
QY 2355 GGCAAAAGAGTCTCTCGAAACATATCTGCTGATGATGAGGCTCGAAGCGGTAA-TA 2413
DB 392 GGCAGAAAGTCTCTCGAAACATATCTGCTGATGATGAGGCTCGAAGCGGTAAATTA 333

QY 2414 AAGACAAAACCGCTGGTGCCTAGCGCTGAAGTCTTACGAGAAAGAGTTGGCGAAAGATC 2473
DB 332 AGGCCAACCCGTTGGTGGCTAGCGCTGAAGTCTTACGAGAAAGAGTTGGCGAAAGATC 273
QY 2474 CACGTATTGGCGCCACCATGGAAGAACGCCAGAAAGGTGAATCATGCCGAACATCCCGC 2533
DB 272 CACGTATTGGCGCCACCATGGAAGAACGCCAGAAAGGTGAATCATGCCGAACATCCCGC 213
QY 2534 AGATGTCCGCTTCTTGGTATGCCGTGCTACTCGGGTATCAACGCCCGCCAGCGGTGCTC 2593
DB 212 AGATGTCCGCTTCTTGGTATGCCGTGCTACTCGGGTATCAACGCCCGCCAGCGGTGCTC 153
QY 2594 AGACTGTGATGAAGCCCTGAAAGACGCGCAGACT 2628
DB 152 AGACTGTGATGAAGCCCTGAAAGACGCGCAGACT 118

RESULT 14
AY080094/c
LOCUS
DEFINITION
AY080094
608 bp DNA linear GSS 06-NOV-2002
uncultured marine virus genomic clone SIO51p3D4L, genomic survey
sequence.
AY080094
GSS.
AY080094.1 GI:24745276
KEYWORDS
uncultured marine virus
SOURCE
uncultured marine virus
ORGANISM
Viruses; environmental samples.
REFERENCE
1 (bases 1 to 608)
AUTHORS
Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M.,
Mead, D., Azam, F. and Rohwer, F.
TITLE
Genomic analysis of uncultured marine viral communities
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
MEDLINE
22294988
PUBMED
12384570
COMMENT
Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@sunstroke.sdsu.edu
Class: shotgun.

FEATURES
Location/Qualifiers
1..608
/organism="uncultured marine virus"
/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone_lib="SIO51p3D4L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
community"
/note="Marine viruses were isolated from 200 liters of
surface seawater using a combination of differential
filtration and density-dependent gradient centrifugation.
Linker-amplified shotgun libraries were constructed by
randomly shearing the total marine viral community DNA,
end-repairing, ligating dsDNA linkers to the ends, and
amplifying the fragments using Vent DNA polymerase. The
resulting fragments were ligated into the pSMART vector
and electroporated into MC12 cells (Lucigen; Middleton,
WI)"

ORIGIN
Query Match 17.9%; Score 589.8; DB 8; Length 608;
Best Local Similarity 99.7%; Pred. No. 2.8e-163;
Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 287 CAGTCTGCCCTGACGCGCGTCGCAAAATGTGCGCGCGATTAAATCTCGCGCGGATCA 346
DB 608 CAGTCTGCCCTGACGCGCGCATCGCAAAATGTGCGCGCGATTAAATCTCGCGCGGATCA 549

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell updates/sec

Title: US-09-765-555B-17
Perfect score: 3300
Sequence: 1 ccgacacatgaatggtgc.....acgacgtccggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sv.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	6	AX202429 Sequence
2	3232.8	98.0	3300	6	AX202430 Sequence
3	3215.2	97.4	3300	6	AX202426 Sequence
4	3215.2	97.4	3300	6	AX202427 Sequence
5	3207.2	97.2	3300	6	AX202428 Sequence
6	2705.8	82.0	7373	6	AX284144 Sequence
7	2705.8	82.0	8101	6	AX172306 Sequence
8	2701	81.8	6648	6	AX377531 Sequence
9	2700	81.8	7475	6	A91965 Sequence 1
10	2700	81.8	7475	6	AR031992 Sequence
11	2700	81.8	7475	6	AR207294 Sequence
12	2698	81.8	9191	6	AX377532 Sequence
13	2676.2	81.1	6724	12	AF097412 Expressio
14	2636.8	79.9	6806	6	AX378208 Sequence
15	2631	79.7	6706	12	AF031088 Shuttle v
16	2598.2	78.7	6748	12	AF031813 Expressio
17	1609.4	48.8	5558	6	CQ794769 Sequence
18	1474	44.7	4700	6	CQ846980 Sequence
19	1462.4	44.3	4921	6	AX768201 Sequence

20	1462.4	44.3	4935	6	AX768200	AX768200 Sequence
21	1462.4	44.3	4945	6	AX768202	AX768202 Sequence
22	1462.4	44.3	4951	6	AX768203	AX768203 Sequence
c	23	1432.4	43.4	6755	12	AF140577 Integrati
	24	1417.6	43.0	1511	6	AX431576 Sequence
	25	1417.6	43.0	4297	12	AY219685 Expressio
	26	1417.6	43.0	4359	12	AY219686 Expressio
	27	1417.6	43.0	4419	12	AY219688 Expressio
	28	1417.6	43.0	4855	12	AY219689 Expressio
	29	1417.6	43.0	4968	12	AY219687 Expressio
	30	1417.6	43.0	6956	12	AY219682 Shuttle e
	31	1417.6	43.0	7018	12	AY219683 Shuttle e
	32	1417.6	43.0	7509	12	AY219684 Shuttle e
	33	1400.8	42.4	4566	12	AF177933 Cloning v
	34	1400.8	42.4	4635	12	AF177932 Cloning v
	35	1400.8	42.4	6399	12	AY302759 Expressio
	36	1400.8	42.4	6587	12	AY289542 Expressio
	37	1362	41.3	5201	6	AR493834 Sequence
	38	1362	41.3	5201	6	AR493835 Sequence
	39	1329.2	40.3	4806	12	AF050464 Expressio
c	40	1238.4	37.5	4803	12	AY243506 Cloning v
c	41	1238.4	37.5	5024	6	CQ846977 Sequence
	42	1237.2	37.5	1922	6	AX478528 Sequence
	43	1204	36.5	4506	12	AB116367 Expressio
	44	1204	36.5	4557	6	E13740 DNA sequenc
	45	1204	36.5	4557	6	AR373251 Sequence

ALIGNMENTS

RESULT 1
AX202429
LOCUS AX202429 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 17 from Patent WO0152620.
ACCESSION AX202429
VERSION AX202429.1 GI:15392177
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 17 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m4 and zinc finger protein 2ZFPm4"

ORIGIN	Query Match	100.0%;	Score 3300;	DB 6;	Length 3300;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3300;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	CGACACCATCGAATGGTGCAGAACCTTTTCGCGGTATGGCATGTAGTCCGCGGATGATAGCCCGGAAGAGA	60		
Db	1	CGACACCATCGAATGGTGCAGAACCTTTTCGCGGTATGGCATGTAGTCCGCGGATGATAGCCCGGAAGAGA	60		
Qy	61	GTCAATTCAGGTGCGTGAATGTGAACACGTAACCTTATACGATTCGCAGAGTATGCCG	120		
Db	61	GTCAATTCAGGTGCGTGAATGTGAACACGTAACCTTATACGATTCGCAGAGTATGCCG	120		
Qy	121	GTGTCCTCTTATCAGACCGCTTTCCCGCGGTGGTGAACACGCGCCAGCCACGCTTTCTCGGAAAA	180		
Db	121	GTGTCCTCTTATCAGACCGCTTTCCCGCGGTGGTGAACACGCGCCAGCCACGCTTTCTCGGAAAA	180		
Qy	181	CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240		

Db	181	CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240
Qy	241	AACAACATGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACATGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Qy	301	ACGGCGCGTCCAAATTTGTCCGGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Db	301	ACGGCGCGTCCAAATTTGTCCGGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Qy	361	TGGTGGTGTGCATGGTGAACGAAGCGCGTGCAGGCTGTAAAGCGCGGTGCACAATC	420
Db	361	TGGTGGTGTGCATGGTGAACGAAGCGCGTGCAGGCTGTAAAGCGCGGTGCACAATC	420
Qy	421	TTCTCGCGCAACGGCTCAGTGGGCTGATTAATTAATCTCGCTCGATGACACGAGATGCCA	480
Db	421	TTCTCGCGCAACGGCTCAGTGGGCTGATTAATTAATCTCGCTCGATGACACGAGATGCCA	480
Qy	481	TTGCTGTGGAAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540
Db	481	TTGCTGTGGAAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540
Qy	541	CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGATCGGGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGATCGGGCGTGGAGCATC	600
Qy	601	TGGTCGCAATGGGTTCACCGCAAAATCGGCTGTAGCGGGCCCAATTAAGTTCGTCTCGG	660
Db	601	TGGTCGCAATGGGTTCACCGCAAAATCGGCTGTAGCGGGCCCAATTAAGTTCGTCTCGG	660
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCGGATAG	720
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCGGATAG	720
Qy	721	CGGAAACGGGAAAGCGACATGGAGTGCCTATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAAACGGGAAAGCGACATGGAGTGCCTATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780
Qy	781	ATGAGGGCATGTTCCCACTCGGATGCTGTTGGCCACGATCAGATGGCGCTGGCGCGAA	840
Db	781	ATGAGGGCATGTTCCCACTCGGATGCTGTTGGCCACGATCAGATGGCGCTGGCGCGAA	840
Qy	841	TGCGCGCATTTACCGAGTCCGGGCTGGCGTGTGGTGGGATATCTCGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACCGAGTCCGGGCTGGCGTGTGGTGGGATATCTCGTAGTGGGATACG	900
Qy	901	ACGATACGGAAGACAGCTCATGTTATATCCGCGCGTTAACACCATCAACAGGATTTTC	960
Db	901	ACGATACGGAAGACAGCTCATGTTATATCCGCGCGTTAACACCATCAACAGGATTTTC	960
Qy	961	GCCTGTGGGGCAACACGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGTGGGGCAACACGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTTGCCGTCTCACTGGTGAAGAAAAACACCTCGCGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCCGTCTCACTGGTGAAGAAAAACACCTCGCGCCCAATA	1080
Qy	1081	CGCAAAACCGCTCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCAGCAGGTTT	1140
Db	1081	CGCAAAACCGCTCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCAGCAGGTTT	1140
Qy	1141	CCGAGTCGGAAGCGGGCAGTGAGGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Db	1141	CCGAGTCGGAAGCGGGCAGTGAGGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Qy	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTCGCG	1260
Db	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTCGCG	1260
Qy	1261	TCAGGCAGCCATCGGAAGCTCTGGTATGGCTGTGCAGTCTGTAATCACTGCATAATTCG	1320

Db	1261	TCAGCGAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATACACTGCATAATTCG	1321
Qy	1321	TGTCGCTCAAGCGCACTCCGCTTCCTGGATAATGTTTTTTTGGCCGCACTCATAAACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACTCCGCTTCGGATAATGTTTTTTTGGCCGCACTCATAAACGGTT	1380
Qy	1381	CTGCGAAATATCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
Db	1381	CTGCGAAATATCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
Qy	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCGCTTAAAGTGTGTTTCA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCGCTTAAAGTGTGTTTCA	1500
Qy	1501	GCATTTCCACCAACAAAGGACCATAGATTATGAAATCTGAAAGGTAACCTGGTAAATCTGG	1560
Db	1501	GCATTTCCACCAACAAAGGACCATAGATTATGAAATCTGAAAGGTAACCTGGTAAATCTGG	1560
Qy	1561	ATTAAACGGCGATAAAGGCTATAACCGGCTCTCGCTGAAAGTCCGTAAGAAATTCGAGAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACCGGCTCTCGCTGAAAGTCCGTAAGAAATTCGAGAAGAT	1620
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATTCGGATAAATCTGGAAGAGAAATTCACACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATTCGGATAAATCTGGAAGAGAAATTCACACAGGTT	1680
Qy	1681	CGCGCAACTCGCGATGCGCTGACATTTATCTTCTGGGCACACGACCGCTTGTGGGCTAC	1740
Db	1681	CGCGCAACTCGCGATGCGCTGACATTTATCTTCTGGGCACACGACCGCTTGTGGGCTAC	1740
Qy	1741	GCTCAATCTGCGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800
Db	1741	GCTCAATCTGCGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCGGTACCGTTACACCGCAAGCTGATTCCTTACCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCGGTACCGTTACACCGCAAGCTGATTCCTTACCCGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGTGATTTTATACAAAGATCTGCTCCGAAACCGCCCAAAACCTCGGAA	1920
Db	1861	GAAGCGTTATCGTGATTTTATACAAAGATCTGCTCCGAAACCGCCCAAAACCTCGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTCACTGGCCGCTGATTCGTCTGACGGGGGTTATGCGTTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGGCCGCTGATTCGTCTGACGGGGGTTATGCGTTTCAAG	2040
Qy	2041	TATGAAACCGCAAGTACAGCATTAAGAGCTGGCGGTGATACCGCTGCGCGGAAGCG	2100
Db	2041	TATGAAACCGCAAGTACAGCATTAAGAGCTGGCGGTGATACCGCTGCGCGGAAGCG	2100
Qy	2101	GGTCTGACCTTCTCGTTGACCTGATTTAAACCAACACATGATGAGACACCGGATTAC	2160
Db	2101	GGTCTGACCTTCTCGTTGACCTGATTTAAACCAACACATGATGAGACACCGGATTAC	2160
Qy	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGCTGG	2220
Db	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGCTGG	2220
Qy	2221	GCATGTTCCACATCGACACACGGAAGTGAATTTATGGTGTACGGTACTGCCGACCTTC	2280
Db	2221	GCATGTTCCACATCGACACACGGAAGTGAATTTATGGTGTACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGGTCAAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT	2340
Db	2281	AAGGGTCAAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAACCTATCTGCTGACTGATGAAGGTCG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAACCTATCTGCTGACTGATGAAGGTCG	2400

QY 841 TCCGCCCAATTACCGAGTCCGGCTCGCGTGTGGTCCGATATCTCGGTAGTGGGATACG 900
Db 841 TCCGCCCAATTACCGAGTCCGGCTCGCGTGTGGTCCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGGTAAACCAACATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGGTAAACCAACATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAACACAGCTGGAGCCGCTTCTGCTCAACTCTCTCAGGCGCAGGCGTGA 1020
Db 961 GCCTGCTGGGGCAACACAGCTGGAGCCGCTTCTGCTCAACTCTCTCAGGCGCAGGCGTGA 1020
QY 1021 AGGCAATCAGCTGTGTCCCGCTTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGTCCCGCTTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAATGAGTGTAGCTCACTCATTTAG 1140
Db 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAATGAGTGTAGCTCACTCATTTAG 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGACGCTGCAATTAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGACGCTGCAATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCGATTAATGTTTTTTCGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCGATTAATGTTTTTTCGCGGACATCAATACGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTGGAATTAATCATCGGCTCGTATTAATGTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTGGAATTAATCATCGGCTCGTATTAATGTGGA 1440
QY 1441 ATTGTAGCGGATTAACATTTTCACAGGAAACAGCGAGTCCGTTAGGTGTTTTTCAG 1500
Db 1441 ATTGTAGCGGATTAACATTTTCACAGGAAACAGCGAGTCCGTTAGGTGTTTTTCAG 1500
QY 1501 GCATTCACCAAGGACCATAGATTATGAAACTGAAGAGTAAACTGGTAACTCTGG 1560
Db 1501 GCATTCACCAAGGACCATAGATTATGAAACTGAAGAGTAAACTGGTAACTCTGG 1560
QY 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAACTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAACTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 CGCGCAACTGCGCATGCGCTGACATTTATCTTGGGCAACAGACCGCTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGCGCATGCGCTGACATTTATCTTGGGCAACAGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGCAAGCTGATTTACCGGATCGCTGTT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGCAAGCTGATTTACCGGATCGCTGTT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTCAACGGCAAGCTGATTCCTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTCAACGGCAAGCTGATTCCTACCGGATCGCTGTT 1860
QY 1861 GAACGCTTATCGCTGATTTATCAAAAGATCTGTCGCGAACCAGCAACCTGGGAA 1920
Db 1861 GAACGCTTATCGCTGATTTATCAAAAGATCTGTCGCGAACCAGCAACCTGGGAA 1920

QY 1921 GAGATCCGCGCTGGATTAAGAACTCAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGGATTAAGAACTCAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTCCTGCTCAACGGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTCCTGCTCAACGGGGGTATGCGTTCAAG 2040
QY 2041 TATGAAAAACCGCAAGTACGACATTAAGAGAGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACCGCAAGTACGACATTAAGAGAGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGGTGGCTGATTAAGAAACAAACACATGATGAGACACCGATTC 2160
Db 2101 GGTCTGACCTTCTGCTGGTGGCTGATTAAGAAACAAACACATGATGAGACACCGATTC 2160
QY 2161 TCCATCGAGAGCTGCTTTTAAAGAGCGAAACAGCGATGACCATCAACGGCGCGCTGG 2220
Db 2161 TCCATCGAGAGCTGCTTTTAAAGAGCGAAACAGCGATGACCATCAACGGCGCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGCTCAACATCCAAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCGCAGT 2340
Db 2281 AAGGCTCAACATCCAAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCGCAGT 2340
QY 2341 CCAGAACAAAGCTGGCAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCAGAACAAAGCTGGCAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATCGAAAAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATCGAAAAACCGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCAGAACATCCCGCAGATGCTCGTTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCAGAACATCCCGCAGATGCTCGTTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCAGCGGCTGCTCAGACTGTCGATGAAGCCCTGAAAGAGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCAGCGGCTGCTCAGACTGTCGATGAAGCCCTGAAAGAGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAAAGGATTTTCAGAAATTC 2700
QY 2701 GGTCTCTCTCTCTGCGCCAGCGGCTCTGAGCCCGGGGAGAGCCCTTATGCTTGT 2760
Db 2701 GGTCTCTCTCTCTGCGCCAGCGGCTCTGAGCCCGGGGAGAGCCCTTATGCTTGT 2760
QY 2761 CCGGAATGTGGTAAGTTCCTTCAGCCAGAGCAGCTCCCTGCTGCGCCACAGGCTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTTCCTTCAGCCAGAGCAGCTCCCTGCTGCGCCACAGGCTACCCAC 2820
QY 2821 ACGGGTGAAGAACCGGTATTAATGCGAGAGTCCAGAAATCTTTTAGCCAGAGCAGCAGC 2880
Db 2821 ACGGGTGAAGAACCGGTATTAATGCGAGAGTCCAGAAATCTTTTAGCCAGAGCAGCAGC 2880
QY 2881 CTGGTGGCCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGATGTGGC 2940
QY 2941 AAGTCTTTTTCAGTGAATGCTGATCTTTCGAGGACCAACCGTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTTTCAGTGAATGCTGATCTTTCGAGGACCAACCGTACTCACACCGGGGAGAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTCTCTCAGAGCTCTCACCTGGTGGCCAC 3060

```

Db 3001 |||||CCCTATGCTTGTCCGGAATGTGGTAAGTCCCTTTACGACACAGTGGTCTCTTGGTTAGACAC 3060
Qy 3061 CAGCGTACCCACACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGGCAAAATCTTTTAGC 3120
Qy 3121 CGCAGCGATACTCGTGGGCCATCAACGCACTCATCTACTGGCGAGAACCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTGGAAAGCCCATCAACGCACTCATCTACTGGCGAGAACCATACAAATGT 3180
Qy 3181 CCAGAATGTGGCAAGTCTTTCTCAACTTCAGSCCAATTTGGTCGTCACCAACGTAACAC 3240
Db 3181 CCAGAATGTGGCAAGTCTTTCTCAACTTCAGSCCAATTTGGTCGTCACCAACGTAACAC 3240
Qy 3241 ACCCGTAAAAAACTAGTGGCCAGCGCCGCGCAGTACCGGTACGAGTTCGCGACTACGCT 3300
Db 3241 ACCCGTAAAAAACTAGTGGCCAGCGCCGCGCAGTACCGGTACGAGTTCGCGACTACGCT 3300

RESULT 3
AX202426 3300 bp DNA linear PAT 30-AUG-2001
LOCUS Sequence 14 from Patent WO0152620.
ACCESSION AX202426
VERSION AX202426.1 GI:15392173
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 14 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m1 and zinc finger protein
ZFPml"

ORIGIN
Query Match 97.4%; Score 3215.2; DB 6; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTACGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCG 120
Db 61 GTCAATTACGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCTCCGGTGTGAAACAGGCCAGCCAGCTTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCTCCGGTGTGAAACAGGCCAGCCAGCTTTCTCGGAAAA 180
Qy 181 CGCGGAAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAAAAGTGTGCTGATTTGGGTGGCGACCTCCAGCTGGCGCTGC 300
Db 241 AACAACTGGCGGGCAAAAGTGTGCTGATTTGGGTGGCGACCTCCAGCTGGCGCTGC 300
Qy 301 ACGCGCGTGCAGAAATGTGCGGGCGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360
Db 301 ACGCGCGTGCAGAAATGTGCGGGCGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360
```

```

Qy 361 TGGTGGTGTGATCGTAGAACGAAAGCGCGCTGGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTGATCGTAGAACGAAAGCGCGCTGGAAGCCTGTAAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACAGCGTCAAGTGGGCTGATCAATACTATCCGCTGGATCACACAGGATGCCA 480
Db 421 TTCTCGCGCAACAGCGTCAAGTGGGCTGATCAATACTATCCGCTGGATCACACAGGATGCCA 480
Qy 481 TTGCTGTGGAAAGTGCCTGCACTAATGTTCCGGGGTATTTCTTGATGTCTCTGACGAGA 540
Db 481 TTGCTGTGGAAAGTGCCTGCACTAATGTTCCGGGGTATTTCTTGATGTCTCTGACGAGA 540
Qy 541 CACCCATCAACAGATATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGATATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGGCCATTAAAGTGTCTGCTCGG 660
Db 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGGCCATTAAAGTGTCTGCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGACTGGAGTGGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTCGCATGTGTTGTCACACGATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTCGCATGTGTTGTCACACGATCAGATGCGCTGGCGGCA 840
Qy 841 TGGCGCCATTACCGAGTCCGGGTGCGGTGTCGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGCCATTACCGAGTCCGGGTGCGGTGTCGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGTCCTGCTCACTGTTGAAAGAAAAACCAACCTGCGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGTCCTGCTCACTGTTGAAAGAAAAACCAACCTGCGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGTGGCACAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGTGGCACAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG 1200
Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATTCATGTCATTAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATTCATGTCATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATTAATGTTTTTTCGCCGAGCATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATTAATGTTTTTTCGCCGAGCATCAACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCACAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCACAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTTCACAGAAACAGCCAGTCCGTTTAGGTGTTTTTCACGA 1500
```


ORIGIN	ZFPm2"	Db	961
Query Match Best Local Similarity 98.4%; Score 3215.2; DB 6; Length 3300; Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;			
Qy	1 CCGACACATCGAATGTTGCAAAACCTTTTCGCGGTATGGCATGATAGCCCGGAAGAGA 60	Qy	AGGCAATCAGCTGTTTGGCCCGTCTCACTGGTGAAGAAAAAACCCACCTCTGGCGCCCAATA 1080
Db	1 CCGACACATCGAATGTTGCAAAACCTTTTCGCGGTATGGCATGATAGCCCGGAAGAGA 60	Db	AGGCAATCAGCTGTTTGGCCCGTCTCACTGGTGAAGAAAAAACCCACCTCTGGCGCCCAATA 1080
Qy	61 GTCAATTCAGGTTGGTGAATGTGAACCAAGTAAGCTTATACGATGTCGCAGAGTATGCCG 120	Qy	CGAAACCCGCTCTCCCGCGGTTTGGCCGATTCATTATGAGCTGGCACACAGAGTTT 1140
Db	61 GTCAATTCAGGTTGGTGAATGTGAACCAAGTAAGCTTATACGATGTCGCAGAGTATGCCG 120	Db	CGAAACCCGCTCTCCCGCGGTTTGGCCGATTCATTATGAGCTGGCACACAGAGTTT 1140
Qy	121 GTGTCTCTTATCAGACCGTTTCCCGCTGGTGAACCAAGGCAGCCAGCTTCTTCGCAAAA 180	Qy	CCGACTCGAAACCGGCGAGTCGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCAATAG 1200
Db	121 GTGTCTCTTATCAGACCGTTTCCCGCTGGTGAACCAAGGCAGCCAGCTTCTTCGCAAAA 180	Db	CCGACTCGAAACCGGCGAGTCGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCAATAG 1200
Qy	181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCACCAACCGCGTGGCAC 240	Qy	GCACAAATTCATGTTTGCAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db	181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCACCAACCGCGTGGCAC 240	Db	GCACAAATTCATGTTTGCAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Qy	241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300	Qy	TCAGGCAGCCATCGGAAGCTGTGTTAGCTGTGCAGTGTGTAATCACTGCATTAATTCG 1320
Db	241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300	Db	TCAGGCAGCCATCGGAAGCTGTGTTAGCTGTGCAGTGTGTAATCACTGCATTAATTCG 1320
Qy	301 ACGCGCGTCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360	Qy	TGTCGCTCAAGGGCACTCCCGTCTGGATAATGTTTTTGGCGCGACATCATACGGTT 1380
Db	301 ACGCGCGTCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360	Db	TGTCGCTCAAGGGCACTCCCGTCTGGATAATGTTTTTGGCGCGACATCATACGGTT 1380
Qy	361 TGGTGGTTCGATGTTGAAGCAAGCGCGCTGGAAGCTCTGTAAGCGCGGTGCACAATC 420	Qy	CTGGCAAAATTTCTGAATGAGCTGTGTGACAAATTAATCATCGCTCGTAAATCTGCTGGA 1440
Db	361 TGGTGGTTCGATGTTGAAGCAAGCGCGCTGGAAGCTCTGTAAGCGCGGTGCACAATC 420	Db	CTGGCAAAATTTCTGAATGAGCTGTGTGACAAATTAATCATCGCTCGTAAATCTGCTGGA 1440
Qy	421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTATCGCTGGATGACCAAGATCCA 480	Qy	GCACTTTCAACCAAGGACCATAGATTATGAAAACCTGGAAGAGGTAAACTGCTAAATCTCG 1560
Db	421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTATCGCTGGATGACCAAGATCCA 480	Db	GCACTTTCAACCAAGGACCATAGATTATGAAAACCTGGAAGAGGTAAACTGCTAAATCTCG 1560
Qy	481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCA 540	Qy	ATTAAACCGGATTAAGGCTTAAACCGCTCTCGCTGAAAGTAAAGAAAATTCGAGAAAGAT 1620
Db	481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCA 540	Db	ATTAAACCGGATTAAGGCTTAAACCGCTCTCGCTGAAAGTAAAGAAAATTCGAGAAAGAT 1620
Qy	541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGGGTGGAGCATC 600	Qy	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAAGAGAAAATTCACACAGGTT 1680
Db	541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGGGTGGAGCATC 600	Db	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAAGAGAAAATTCACACAGGTT 1680
Qy	601 TGGTCGATTTGGGTCAACAGCAAAATCGCGCTGTTTGGCGGCGCAATTAAGTTCTGTTCTCG 660	Qy	GGGCAACTCTGGCGATGGCCCTGACATTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Db	601 TGGTCGATTTGGGTCAACAGCAAAATCGCGCTGTTTGGCGGCGCAATTAAGTTCTGTTCTCG 660	Db	GGGCAACTCTGGCGATGGCCCTGACATTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Qy	661 CGCGTCTGCGTCTGGCTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720	Qy	GCTCAATCTGGGCTGTTGGCTGAAATCACCCGGCAAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db	661 CGCGTCTGCGTCTGGCTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720	Db	GCTCAATCTGGGCTGTTGGCTGAAATCACCCGGCAAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy	721 CGGAAACGGGAAGGCACTGGAGTGCCATGTCGGGTTTTCAACAAACCATGCAATGCTGA 780	Qy	CCGTTTACCTGGGATCCCGTACGTTTCAACCGGCAAGCTGATTTACCCGATTCGCTGTT 1860
Db	721 CGGAAACGGGAAGGCACTGGAGTGCCATGTCGGGTTTTCAACAAACCATGCAATGCTGA 780	Db	CCGTTTACCTGGGATCCCGTACGTTTCAACCGGCAAGCTGATTTACCCGATTCGCTGTT 1860
Qy	781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGGCGCTGGCGCAA 840	Qy	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCGCAAAACCTCGGAA 1920
Db	781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGGCGCTGGCGCAA 840	Db	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCGCAAAACCTCGGAA 1920
Qy	841 TGCGGCGCATTTACGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900	Qy	GAGATCCCGCGCTGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Db	841 TGCGGCGCATTTACGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900	Db	GAGATCCCGCGCTGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Qy	901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTAAACCAACCATCAAAACGATTTTC 960	Qy	CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db	901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTAAACCAACCATCAAAACGATTTTC 960	Db	CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy	961 GCCTGCTGGGCAAAACCAAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020	Qy	TATGAAAAACCGCAAGTACGACATTAAAGACGTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db		Db	TATGAAAAACCGCAAGTACGACATTAAAGACGTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100

QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATTAAGCGCAAAACAGCGATGACCATCAACGGCCCGGTG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATTAAGCGCAAAACAGCGATGACCATCAACGGCCCGGTG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGGTAAACGGTACTCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGGTAAACGGTACTCGGACCTTC 2280
QY 2281 AAGGTCACCAATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAAGCGCCGCCAGT 2340
Db 2281 AAGGTCACCAATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAAGCGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATCAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATCAAGGTCTG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCGTGGTCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGTGGTCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTGCGGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTGCGGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGAGATGTCGCGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGCGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAATTC 2700
QY 2701 GGATCTCTCTCTGTCGCGCCAGCGGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTCTCTGTCGCGCCAGCGGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAAGTCTCTCAGCCAGAGCAGCTCCCTGGTGGCCACACGAGTACCCAC 2820
Db 2761 CCGGAATGTGTAAAGTCTCTCAGAGCTCTCACTGGTGGCCACACGAGTACCCAC 2820
QY 2821 ACGGTTGAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTTGAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCATCAACGACCTACTTGGCGAGAGCCATACAACTCTCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGACCTACTTGGCGAGAGCCATACAACTCTCAGATGTGGC 2940
QY 2941 AAGTCTTTCAGTGTATGTCGTGATCTTTGCGAGGACCAACGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCAGTGTATGTCGTGATCTTTGCGAGGACCAACGTACTCACACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACTGGTGGCCAC 3060
QY 3061 CAGGTTACCCACACGGGTGAAAAACCGTATAAATGCCCAGAGTGGCAAACTTTTAGC 3120
Db 3061 CAGGTTACCCACACGGGTGAAAAACCGTATAAATGCCCAGAGTGGCAAACTTTTAGC 3120
QY 3121 CGCAGCGATAACTGGTGGCGCCATCAACGCACTCATATCGCGAGAGAGCCATACAAATGT 3180
Db 3121 CAGGCCGGCCACTGGCGAGCCATCAACGCACTCATATCGCGAGAGAGCCATACAAATGT 3180

QY 3181 CCAGAAATGTGCAAGTCTTTCTCAACTTCAGGCAATTTGGTCGTCACCAACGTACTCAC 3240
Db 3181 CCAGAAATGTGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCGGCACCACGTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGCGAGTACCGGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGCGGCGGTAACCGTACGACGTTCCGGACTACGCT 3300
RESULT 5
AX202428 3300 bp DNA linear PAT 30-AUG-2001
LOCUS Sequence 16 from Patent WO0152620.
DEFINITION AX202428
ACCESSION AX202428
VERSION AX202428.1 GI:15392176
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 16 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES Location/Qualifiers
source 1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m3 and zinc finger protein ZFPm3"

ORIGIN
Query Match 97.2%; Score 3207.2; DB 6; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 1 CCAGACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGAAGAGA 60
Db 1 CCAGACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGAAGAGA 60
QY 61 GTCAAATTCAGGCTGGTGAATGTGAAACCAAGTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAAATTCAGGCTGGTGAATGTGAAACCAAGTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACGCTTTCCCGGTGTGAACCAAGCGACCGACGTTTCTGCGAAA 180
Db 121 GTGTCTCTTATCAGACGCTTTCCCGGTGTGAACCAAGCGACCGACGTTTCTGCGAAA 180
QY 181 CCGCGGAAAAGTGGAAAGCGGCGATGCGGAGCTGAATTTACATTTCCCAACCGGTCGCAC 240
Db 181 CCGCGGAAAAGTGGAAAGCGGCGATGCGGAGCTGAATTTACATTTCCCAACCGGTCGCAC 240
QY 241 AACAACTGGCGGCAACACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCGCAAAATTTGTCGGCGGATTAATTTCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTCGCAAAATTTGTCGGCGGATTAATTTCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGTCGATGGTAGAACGAAGCGGCTGGAAGCGGCTGTAAGCGGCGGTGCAAAATC 420
Db 361 TGGTGTGTGTCGATGGTAGAACGAAGCGGCTGGAAGCGGCTGTAAGCGGCGGTGCAAAATC 420
QY 421 TTCTCCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGTCGATGACCGAGTGCACA 480
Db 421 TTCTCCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGTCGATGACCGAGTGCACA 480
QY 481 TTGCTGTGGAAGCTGCTGCATTAATTTCCGCGCTTATTTCTTGATGTCTCTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCATTAATTTCCGCGCTTATTTCTTGATGTCTCTGACACAGA 540

QY 541 CACCCATCAAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCATTGGGTCACACGAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTCACACGAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 840
Db 781 ATGAGGGCATCGTTCCCACTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 840
QY 841 TCGCGGCCATTACCGAGTCCGGCTCGCGGTTGGTGCAGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGGCCATTACCGAGTCCGGCTCGCGGTTGGTGCAGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAATCAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAATCAAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGAGTGGGACCGCTTCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGTGGGACCGCTTCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 1080
QY 1081 CGCAACCGCCTCTCCCGCGGTTGGCGATTCATTAAATGCAAGTGGCAACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGATTCATTAAATGCAAGTGGCAACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTCCCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTCCCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTGTGA 1440
Db 1381 CTGGCAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTGTGA 1440
QY 1441 ATTGTAGCGGATAACAATTTCAACAGGAACAGCCAGTCCGTTTAGGTGTTTCAAGA 1500
Db 1441 ATTGTAGCGGATAACAATTTCAACAGGAACAGCCAGTCCGTTTAGGTGTTTCAAGA 1500
QY 1501 GCATTTCAACACAGGACCATAGATTATGAAGAACTGAAGAGGTAAACCTGTAATCTGG 1560
Db 1501 GCATTTCAACACAGGACCATAGATTATGAAGAACTGAAGAGGTAAACCTGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAAAGGCTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATATAAATGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATATAAATGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGGCAACTGCGGATGGCCCTGACATATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGGATGGCCCTGACATATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGACAAGCGTTCCAGGAACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGACAAGCGTTCCAGGAACAAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATGGCTAGCTTACAACGGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CGGTTTACCTGGGATGGCTAGCTTACAACGGCAAGCTGATGCTTACCCGATCGCTGT 1860
QY 1861 GAAGGTTTATCGCTGATTATAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTATAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAACACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAACACATGAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGAAGAGCTGCGCTTTAATAAGCGAAGACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGAAGAGCTGCGCTTTAATAAGCGAAGACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACGACGAAAGTGAATTAATGTTGTAACGCTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACGACGAAAGTGAATTAATGTTGTAACGCTGCGGACCTTC 2280
QY 2281 AAGGTCACACCATCCAAACCGTTCTGCGCTGCTGAGCGCAGGTATTAAACGCCCACT 2340
Db 2281 AAGGTCACACCATCCAAACCGTTCTGCGCTGCTGAGCGCAGGTATTAAACGCCCACT 2340
QY 2341 CGAAACAAAGAGCTGGCAAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CGAAACAAAGAGCTGGCAAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGCCTGAGCGCTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGCCTGAGCGCTGAAGTCTTTACGAGGAAG 2460
QY 2461 TTGGCAAGAGTCCAACGTTATGCGCCCACTGGAACAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCAAGAGTCCAACGTTATGCGCCCACTGGAACAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACAACTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACAACTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGTTGGCCCGACGCGGCCCTTCGAGCCCGGGGAGAACCCCTATGCTGT 2760


```
Db 2701 GGAATCCTCTCTCTGTGGCCAGCGCCCTCGAGCCGGGAGAACCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTTTCAGCAGAGCAGCTCCCTGGTGGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTTCAGCAGATCCTGGCCACCTGGTTCGGCCACAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTTAGCCAGACGACGACG 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTTAGCCAGACGCGCTCC 2880
Qy 2881 CTGGTGGCCCATCAGCAGCACTCATACTGGCAGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATACTGGCAGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Qy 2941 AAGTCTCTTCAGTGAATGTGCTGATCTTGGAGGACCAACAGTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTCTTCAGCAGAGCTCCAGCTGTGGCCACCAAGTACTCACACCGGGGAGAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCAGAGCTCTCACCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCAGCAGAGCAGCTCCCTGGTGGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTAGT 3120
Qy 3121 CGCAGCGATAACCTGGTGGCCATCAACGCACTCATACTGGCAGAGAGCCATACAAATGT 3180
Db 3121 GACTGCCGCGACCTTGTCCGCATCAACGCACTCATACTGGCAGAGAGCCATACAAATGT 3180
Qy 3181 CCAGATGTGGCAAGTCTTCTCACTTCAGGCCATTTGGTCGCTACCAACGTAAGTAC 3240
Db 3181 CCAGATGTGGCAAGTCTTCTCCCAATCCAGCCATCTCGTCCGCGACCAACGTAAGTAC 3240
Qy 3241 ACCGCTAAAAAATAGTGGCCAGCGCCGAGTACCGTACGAGCTTCGGAGTACGCT 3300
Db 3241 ACCGCTAAAAAATAGTGGCCAGCGCCGAGTACCGTACGAGCTTCGGAGTACGCT 3300

RESULT 6
AX284144 7373 bp DNA linear PAT 20-NOV-2001
LOCUS
DEFINITION Sequence 1 from Patent WO0178756.
ACCESSION AX284144
VERSION AX284144.1 GI:17044832
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1. Wiedersanders, B. and Maubach, G.
AUTHORS Agent for postoperative use after the removal of bone tumours
TITLE Patent: WO 0178756-A 1 25-OCT-2001;
JOURNAL Dupuy Biotech Jena GmbH (DE)
FEATURES
Location/Qualifiers
source
1..7373
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Expressionsvektor pMalc2cchbnp2"
1528..3453
/note="unnamed protein product; /produkt=Fusionsprotein
MBP Cystatin C BMP-2"
/codon_start=1
/transl_table=11
/protein_id="CAD12160.1"
/db_xref="GI:17044833"
/translation="MKIEEKLVIVINGDKYGNLAEBVGKPKFDGKIVTVHPDKL
EEKFPQVAATGDGPDIIIFWADHFGGTAQSGLLAEITPDRAFDQKLYPFTWDAVRNG
KLIAYPEAVSALSYKNDLLPNPPTKWEIIPALDKELKAGKSAFMNLQEPYFTWP
LIAADGGYAVENGYKIDKVDGNAGAGLTFILVDLKNKHMNADTDSIAEAAFP
NKGETAMTINGPMAWSNIDTSKVNYGVTLPTFKGQPSKFPVGLSAGINAASPKNEL
```

```
misc_feature
2689
/note="Faktor Xa Protease /position=388 (AMINO
ACIDsequenz)"
3067..3108
/misc_feature
/note="spacer zwischen Cystatin C und BMP-2"
ORIGIN
Query Match 82.0%; Score 2705.8; DB 6; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGCTGCAAAACCTTTCGCGGTATGGCATAGTAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGCTGCAAAACCTTTCGCGGTATGGCATAGTAGCGCCCGGAAGAGA 60
Qy 61 GTCMAATTCAGGTTGGTGAATGTGAACACCACTAAGTTATACGATGTGCGAGAGTATCGCG 120
Db 61 GTCMAATTCAGGTTGGTGAATGTGAACACCACTAAGTTATACGATGTGCGAGAGTATCGCG 120
Qy 121 GTGTCCTTATCAGACCGTTTCCGCGGTGTGAACAGGCGCAGCAGCTTCTTCGCAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCGCGGTGTGAACAGGCGCAGCAGCTTCTTCGCAAAA 180
Qy 181 CGCGGAAAAAAGTGGAGCGCGATGGCGGAGCTGAATTAACATTCACCAACCGGTTGGCAC 240
Db 181 CGCGGAAAAAAGTGGAGCGCGATGGCGGAGCTGAATTAACATTCACCAACCGGTTGGCAC 240
Qy 241 AACCACTGGCGGGGCAACAGTCTGTTGCTGATTTGGCGTTGCGCACTCCAGTCTGGCCCTGC 300
Db 241 AACCACTGGCGGGGCAACAGTCTGTTGCTGATTTGGCGTTGCGCACTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGCTCGCAAAATTTGTCGCGCGATTAATAATCTCGCGCGGATCACTCGGTTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGCGCGATTAATAATCTCGCGCGGATCACTCGGTTGCCAGCG 360
Qy 361 TGTGTTGTCGATGGTAGAACGAGCGCGCTCGAAGCCTGTAAAGCGCGCGGTGCACAATC 420
Db 361 TGTGTTGTCGATGGTAGAACGAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAACATCTCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAACATCTCGCTGGATGACCAAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Qy 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGTGTCGATTTGGGTACCGCAAAATCGCTGTTAGCGGGCCATTAAGTCTGCTCTCG 660
Db 601 TGTGTCGATTTGGGTACCGCAAAATCGCTGTTAGCGGGCCATTAAGTCTGCTCTCG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAACGGCACTGGAGTGGCATGTCCGGTTTTCAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAACGGCACTGGAGTGGCATGTCCGGTTTTCAAACCAATGCTGA 780
Qy 781 ATGAGGGCATCGTTTCCCACTGCTGCTGTTGCCAAACCATCAGATGGCGCTGGGGCGCA 840
Db 781 ATGAGGGCATCGTTTCCCACTGCTGCTGTTGCCAAACCATCAGATGGCGCTGGGGCGCA 840
```


QY 841 TCGCGCCATTACCGAGTCCGGGCTCGCGTTGGTGGGATATCTCGGTAGTGGATACG 900
DB 841 TCGCGCCATTACCGAGTCCGGGCTCGCGTTGGTGGGATATCTCGGTAGTGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACACCACTCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACACCACTCAACAGGATTTTC 960
QY 961 GCTGCTGGGGCAACACAGCGTGGACCGCTTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
DB 961 GCTGCTGGGGCAACACAGCGTGGACCGCTTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCCAATA 1080
QY 1081 CCGAAACCGCTCTCCCGCGGTTGGCGGATTCATTATGAGCTGGCGACCAAGGTTT 1140
DB 1081 CCGAAACCGCTCTCCCGCGGTTGGCGGATTCATTATGAGCTGGCGACCAAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTATGTTGACGCTTATCATGCACTGCAGGTCGCAATGCTTCTGGCG 1260
DB 1201 GCACAATTTCTATGTTGACGCTTATCATGCACTGCAGGTCGCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGCATTCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320
DB 1261 TCAGGCGCATTCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAACGGTT 1380
DB 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATGAGCTGTTGACATTAATCATCGCTCGTATAATGTGTGA 1440
DB 1381 CTGGCAATATTTCTGAATGAGCTGTTGACATTAATCATCGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTAGCGGATAACAATTTCCACAGGAACAGCCAGTCGTTTAGTGTTTTTCAGA 1500
DB 1441 ATTGTAGCGGATAACAATTTCCACAGGAACAGCCAGTCGTTTAGTGTTTTTCAGA 1500
QY 1501 GCACCTTCAACAAGGACCATAGATTATGAANAATCGAAGAGTTAACTGGTAATCTGG 1560
DB 1501 GCACCTTCAACAAGGACCATAGATTATGAANAATCGAAGAGTTAACTGGTAATCTGG 1560
QY 1561 ATTACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
DB 1561 ATTACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACAGCCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCCCGGCAACAGGTTTCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCCCGGCAACAGGTTTCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCCGTAGCTTACACGCGAAGCTGATGCTTACCCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGATGCCGTAGCTTACACGCGAAGCTGATGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGTGAATTTATACAAAGATCTGTCGCGAAACCGGCCAAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGTGAATTTATACAAAGATCTGTCGCGAAACCGGCCAAAACCTGGGAA 1920

QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTCTTCACTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTCTTCACTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGAAGTACGACATTAAGAACGTTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
DB 2041 TATGAAAAACGGAAGTACGACATTAAGAACGTTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACACATGAATGCGAGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACACATGAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGAAACAGCGATGACCATCAAGCGGCCGCTGG 2220
DB 2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGAAACAGCGATGACCATCAAGCGGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACACAGCAAAGTGAATTAATGTGTAAACGGTACTGCGCACTTC 2280
DB 2221 GCATGGTCCAAATCGACACACAGCAAAGTGAATTAATGTGTAAACGGTACTGCGCACTTC 2280
QY 2281 AAGGTCACACCATCCAAACCGTTGGTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT 2340
DB 2281 AAGGTCACACCATCCAAACCGTTGGTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAACACTATCTGCTGATGAAGTCTG 2400
DB 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAACACTATCTGCTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGTGGTGGTGAAGTCTTACGAGAAAG 2460
DB 2401 GAAGCGGTTAATAAAGACAAACCGCTGGTGGTGGTGAAGTCTTACGAGAAAG 2460
QY 2461 TTGGGAAAGATCCACGATTTGCGCCACCATGGAACGCCAGAAAGGTTGAATCATG 2520
DB 2461 TTGGGAAAGATCCACGATTTGCGCCACCATGGAACGCCAGAAAGGTTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC 2580
DB 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTG 2640
DB 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTG 2640
QY 2641 AACACAAACAATTAACAATAACAACCTCGGGATCGAGGAAAGGATTTCAAGATTC 2700
DB 2641 AACACAAACAATTAACAATAACAACCTCGGGATCGAGGAAAGGATTTCAAGATTC 2700
QY 2701 GGATCCTCTCTCTGTCGGCCAGCGCCCTTCGAGCCCGG 2741
DB 2701 GGATCCTCTCTCTGTCGGCCAGCGCCCTTCGTCGGAGG 2741

RESULT 7

AXI172306
LOCUS AXI172306 8101 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 5 from Patent WO014467.
ACCESSION AXI172306
VERSION AXI172306.1 GI:14597487
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Goubin-Gramatica,F., Ducommun,B. and Prevost,G.
TITLE Method for obtaining human cdc25 phosphatases and method for identifying human cdc25 phosphatase modulators
JOURNAL Patent: WO 014467-A 5 21-JUN-2001;
SOCIETE DE CONSEILS DE RECHERCHES ET D'APPLICATIONS SCIENTIFIQUES

FEATURES		Location/Qualifiers	ORIGIN
source			
1. .8101			
/organism="synthetic construct"			
/mol_type="unassigned DNA"			
/db_xref="taxon:32630"			
/note="souche JM109 / pMAL - Hs Cdc25C"			
Query Match		82.0%; Score 2705.8; DB 6; Length 8101;	
Best Local Similarity		99.9%; Pred. No. 0;	
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	CGACACCATGAAATGGTGAACAACTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
DB	1	CGACACCATGAAATGGTGAACAACTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCCAGAGTATGCCG	120
DB	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCCAGAGTATGCCG	120
QY	121	GTGTCCTCTTATCAGACCCGTTTCCCGGTGGTGAACCAAGCCAGCCACGCTTTCTGCGAAA	180
DB	121	GTGTCCTCTTATCAGACCCGTTTCCCGGTGGTGAACCAAGCCAGCCACGCTTTCTGCGAAA	180
QY	181	CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC	240
QY	241	AACAACTCGCGGCAAAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTCGCGGCAAAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGCTCGCAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCGACGC	360
DB	301	ACGCGCGCTCGCAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCGACGC	360
QY	361	TGGTGGTTCGATGTGAAGCAAGCGCGCTCGAAGCGCTGTAAGCGCGCGGTGCACAATC	420
DB	361	TGGTGGTTCGATGTGAAGCAAGCGCGCGTGAAGCGCGCTGTAAGCGCGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACGGTCACTGGGTGATCAATTAATCTCGCTCGATCACTGGGTGCGACGC	480
DB	421	TTCTCGCGCAACGGTCACTGGGTGATCAATTAATCTCGCTCGATCACTGGGTGCGACGC	480
QY	481	TTGCTGTGGAAGCTGCCTGCACATTAATGTTCCGGCGTTATTTCTTGAATGTTCTGACCCAGA	540
DB	481	TTGCTGTGGAAGCTGCCTGCACATTAATGTTCCGGCGTTATTTCTTGAATGTTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC	600
QY	601	TGGTCCGATTCGGGTCAACAGCAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCGCG	660
DB	601	TGGTCCGATTCGGGTCAACAGCAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCGCG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAGCGACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA	780
DB	721	CGGAACGGGAGCGACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCGATCGTTCCCACTCGCATGCTGGTTGCAACGATCAGATGGCGCTGGCGCAA	840
DB	781	ATGAGGCGATCGTTCCCACTCGCATGCTGGTTGCAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGCGCATTCACGAGTCCGGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGCGCATTCACGAGTCCGGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900

QY	901	ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCCACCATCAAAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCCACCATCAAAACAGGATTTTC	960
QY	961	GCCTGTGGGGCAAAACCAAGCGGTGGACCGCTTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
DB	961	GCCTGTGGGGCAAAACCAAGCGGTGGACCGCTTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCGCGTCTCACTGTGTAAGAAAAAACCCACCTGGCGGCCAATA	1080
DB	1021	AGGCAATCAGCTGTTGCGCGTCTCACTGTGTAAGAAAAAACCCACCTGGCGGCCAATA	1080
QY	1081	CGCAAAACCGCTCTCTCCCGCGCGTGGCCGATTCATTAAATGACGCTGGCAGCAGAGTTT	1140
DB	1081	CGCAAAACCGCTCTCTCCCGCGCGTGGCCGATTCATTAAATGACGCTGGCAGCAGAGTTT	1140
QY	1141	CCGACCTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTTAG	1200
DB	1141	CCGACCTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTTAG	1200
QY	1201	GCACAAATTCATGTTTTCAGAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATTCATGTTTTCAGAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCCATCGAAAGCTGTGTATGCTGTGCGAGTCTGTAATCACTGCATAATTCG	1320
DB	1261	TCAGGACGCCATCGAAAGCTGTGTATGCTGTGCGAGTCTGTAATCACTGCATAATTCG	1320
QY	1321	TGTCGCTCAAGGGCGACTCCCGTTTGGAATAATGTTTTTTTGGCGCGACATCATACCGTT	1380
DB	1321	TGTCGCTCAAGGGCGACTCCCGTTTGGAATAATGTTTTTTTGGCGCGACATCATACCGTT	1380
QY	1381	CTGCAAAATTTCTGAAATCAGCTGTGTGCAATTAATCACTCGGCTGTAATGTGTGGA	1440
DB	1381	CTGCAAAATTTCTGAAATCAGCTGTGTGCAATTAATCACTCGGCTGTAATGTGTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTATAGTGTGTTTACGA	1500
DB	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTATAGTGTGTTTACGA	1500
QY	1501	GCATTCACCAACAGGACCATGATTAATGAAACTGGAAGAGTAACTGGTAAATCTGG	1560
DB	1501	GCATTCACCAACAGGACCATGATTAATGAAACTGGAAGAGTAACTGGTAAATCTGG	1560
QY	1561	ATTAAACGGCGATAAAGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACGGCGATAAAGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCACACAGTT	1680
DB	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCACACAGTT	1680
QY	1681	CGCGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
DB	1681	CGCGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCGTGTGGCTGAAATCACCCCGCAACAAAGCGTTCCAGGACAAAGCTGAT	1800
DB	1741	GCTCAATCTGGCGTGTGGCTGAAATCACCCCGCAACAAAGCGTTCCAGGACAAAGCTGAT	1800
QY	1801	CCGTTTACCTGGGATGCGGTACGTTTAAACGCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
DB	1801	CCGTTTACCTGGGATGCGGTACGTTTAAACGCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTCCGCAACCCCGCAAAACCTGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTCCGCAACCCCGCAAAACCTGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGGAAAGGTAAAGACGCGCTGATGTTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGGAAAGGTAAAGACGCGCTGATGTTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATGTGCTGTAACGGGGTTATGCTGTTCAAG	2040

Db 1981 CTGCAAGAACCGTACTTCACTTGGCGCTGATTGCTGTGACGGGGTTATCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGATACCGTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGATACCGTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAATAAACAACATGAATGACGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAATAAACAACATGAATGACGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCCCTTAAATAAGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTAAATAAGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACCGAAGAGTGAATTATGGTGTAAACGCTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCGAAGAGTGAATTATGGTGTAAACGCTACTGCCGACCTTC 2280
Qy 2281 AAGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Qy 2341 CGAACAAGAGCTGCGAAGAGTTCTCGAAGAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CGAACAAGAGCTGCGAAGAGTTCTCGAAGAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAAGCGGTTAATAAGACAAACCGTGGTGGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAAGCGGTTAATAAGACAAACCGTGGTGGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGGCCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAATCCCGCAGATGTCGCTTCTGGTATGCGGTACTCGGTGATCAACGCC 2580
Db 2521 CCGAATCCCGCAGATGTCGCTTCTGGTATGCGGTACTCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTCTCGATGAAGCCCTGAAAGACGCCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTCGATGAAGCCCTGAAAGACGCCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAACAACAATAACAAATTAACAACCTCGGATCGAGGAAGGATTTCAAGATTC 2700
Db 2641 AACAACAACAATAACAAATTAACAACCTCGGATCGAGGAAGGATTTCAAGATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
AX377531
LOCUS AX377531 6648 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 8 from Patent WO0212553.
ACCESSION AX377531
VERSION AX377531.1 GI:19573717
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE other sequences; artificial sequences.
1
AUTHORS Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrendsdorf,H. and Muth,J.
TITLE Method for detecting mutations in nucleotide sequences
JOURNAL Patent: WO 0212553-A 8 14-FEB-2002;
Nanogen Recognomics GmbH (DE)
FEATURES
1. .6648
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

/note="Plasmid pMAL-c2x"

ORIGIN	Query Match Best Local Similarity Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1	CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGGCCCGCGAAGAGA 60
Db 1	CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGGCCCGCGAAGAGA 60
Qy 61	GTCRAATTCAGGGTGTGAATGTGAAACCAAGTAAAGCTTATACGATGTCGACAGTATGCG 120
Db 61	GTCRAATTCAGGGTGTGAATGTGAAACCAAGTAAAGCTTATACGATGTCGACAGTATGCG 120
Qy 121	GTGTCCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCACGCTTTCTCGGAAAA 180
Db 121	GTGTCCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCACGCTTTCTCGGAAAA 180
Qy 181	CGCGGAAAAAGTGAAGCGGGATGGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCNC 240
Db 181	CGCGGAAAAAGTGAAGCGGGATGGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCNC 240
Qy 241	AACAACTCGCGGCAAAACAGTCGTTGCTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGC 300
Db 241	AACAACTCGCGGCAAAACAGTCGTTGCTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGC 300
Qy 301	ACGCGCGCTCGCAATTTGTCGCGCGATTAATCTCGCGCCGATCAACTGGGTGGCAGCG 360
Db 301	ACGCGCGCTCGCAATTTGTCGCGCGATTAATCTCGCGCCGATCAACTGGGTGGCAGCG 360
Qy 361	TGGTGGTGTGATGTAAGAACGAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAAATC 420
Db 361	TGGTGGTGTGATGTAAGAACGAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAAATC 420
Qy 421	TTCTCGCCCAACGGCTCAGTGGGCTGATTAATTAATCTCGCTGATGACACAGGATGCCA 480
Db 421	TTCTCGCCCAACGGCTCAGTGGGCTGATTAATTAATCTCGCTGATGACACAGGATGCCA 480
Qy 481	TTGCTGTGGAAGCTCCCTGCATTAATGTTCCGGCGCTTAATTTCTGATGTTCTCTGACCAG 540
Db 481	TTGCTGTGGAAGCTCCCTGCATTAATGTTCCGGCGCTTAATTTCTGATGTTCTCTGACCAG 540
Qy 541	CACCCATCAACAGTATTAATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541	CACCCATCAACAGTATTAATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601	TGCTCGCATTTGGTCAACGAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG 660
Db 601	TGCTCGCATTTGGTCAACGAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG 660
Qy 661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721	CGAAACGGGAAGCGACGTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721	CGAAACGGGAAGCGACGTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781	ATGAGGCGATCGTTCCCACTCGGATGCTGTTGCCAAGATCAGATGCGGCTGGGGCGAA 840
Db 781	ATGAGGCGATCGTTCCCACTCGGATGCTGTTGCCAAGATCAGATGCGGCTGGGGCGAA 840
Qy 841	TGCGCGCATTAACCGAGTCCCGGCTCGCGTGTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841	TGCGCGCATTAACCGAGTCCCGGCTCGCGTGTGGTGGCGATATCTCGGTAGTGGGATAG 900
Qy 901	ACGATACCGAAGACAGCTCATGTTTATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
Db 901	ACGATACCGAAGACAGCTCATGTTTATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
Qy 961	GCCTGCTGGGGCAAAACAGCGTGGAGCCGCTTGCTGCAACTCTCTCAGGCCAGGCGGTGA 1020
Db 961	GCCTGCTGGGGCAAAACAGCGTGGAGCCGCTTGCTGCAACTCTCTCAGGCCAGGCGGTGA 1020

```
Db 961 GCCTGCTGGGCAAAACACGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
Qy 1021 AGGCAATCAGCTGTGTCCTGCTCACTGGTGAAGAAACACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGTCCTGCTCACTGGTGAAGAAACACCCCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTGGCGATTCAATTAATGAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTGGCGATTCAATTAATGAGCTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTCAGCGCAACCAATTAATGAGTACCTACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTCAGCGCAACCAATTAATGAGTACCTACTCATTTAG 1200
Qy 1201 GCACAATTTCTCATGTGTGACAGCTTATCATGACCTGCACGGTGCAACATGCTTGGCG 1260
Db 1201 GCACAATTTCTCATGTGTGACAGCTTATCATGACCTGCACGGTGCAACATGCTTGGCG 1260
Qy 1261 TCAGGACGCATCGGAAGCTGGTATGCTGTGAGGTGCGTAAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGGTATGCTGTGAGGTGCGTAAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTTCGCCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTTCGCCGACATCATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTAAATCGTGTGGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTAAATCGTGTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTTACACAGAAACAGCCAGTCGGTTAGGTGTTTCAGGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGAAACAGCCAGTCGGTTAGGTGTTTCAGGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAGCTGAAGAGTAACTGCTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAGCTGAAGAGTAACTGCTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATAACTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATAACTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGACAAAGGTTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGACAAAGGTTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGTACGTTACACCGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTACACCGCAAGCTGATGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGCAACCCCGCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGCAACCCCGCAAAACCTCGGAA 1920
Qy 1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGGTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGGTATGCTTCAAG 2040
Qy 2041 TATGAAAACCGCAAGTACGACATTAAAGACGTGGGCTGGATACGCTGCGCGAAAGCG 2100
Db 2041 TATGAAAACCGCAAGTACGACATTAAAGACGTGGGCTGGATACGCTGCGCGAAAGCG 2100
```

```
Qy 2101 GGTCTGACCTTCCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCATGACATCAACGCGCCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCATGACATCAACGCGCCCGTGG 2220
Qy 2221 GCATGTGTCACATCGACACACAGCAAAAGTGAATTTATGGTTPAACGGTATGTCGCACTTC 2280
Db 2221 GCATGTGTCACATCGACACACAGCAAAAGTGAATTTATGGTTPAACGGTATGTCGCACTTC 2280
Qy 2281 AAGGCTCAACCATCAACACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGCTCAACCATCAACACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Qy 2401 GAAGCGTTTAAATAAGCAAAACCGCTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGCAAAACCGCTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCACGTATTTGCCGCCACCATGMAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTTGCCGCCACCATGMAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGTGTCCGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGTGTCCGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTGTGAGTGTGATGAAGCCCTGAAGACCGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGAGTGTGATGAAGCCCTGAAGACCGCGAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATAACAAATAACAACTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAACAAATAACAACTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 9
LOCUS A91965 7475 bp DNA circular PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9822496.
ACCESSION A91965
VERSION A91965.1 GI:6740811
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood,M.R. and Hurst,D.N.
TITLE ANTIVIRAL PEPTIDE DERIVATIVES
JOURNAL Patent: WO 9822496-A 1 28-MAY-1998;
HOFFMANN LA ROCHE (CH)
FEATURES
source 1..7475
Location/Qualifiers
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGATGGTGCAAAACCTTTCCGGGTATGCGGTATGCGCGGAGGATTTTCAGAAAGA 60
```

|||||
1 CGGACACCATCGAATGGTGCAGAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTTCGCGAGAGTATGCCG 120
61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTTCGCGAGAGTATGCCG 120
121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAATAA 180
121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAATAA 180
181 CSCGGGAAAAAGTGAAGCGCGATGGGGAGCTGAATTTACATTTCCCAACCCGCTGGGCAC 240
181 CSCGGGAAAAAGTGAAGCGCGATGGGGAGCTGAATTTACATTTCCCAACCCGCTGGGCAC 240
241 AACCAACTCGCGGCGCAACAGTCTGTGCTGATTTGGGGTTGGCCACCTCCAGTCTGGGCCCTGC 300
241 AACCAACTCGCGGCGCAACAGTCTGTGCTGATTTGGGCTTGGCCACCTCCAGTCTGGGCCCTGC 300
301 AGCGCGCTGCGCAAAATTTGTCGGCGCATTTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
301 AGCGCGCTGCGCAAAATTTGTCGGCGCATTTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
361 TGGTGGTGTGCGATGTAGAACGAAGGGCGTCGAAGCCCTGTAAAGCGGCGGTGCAACAATC 420
361 TGGTGGTGTGCGATGTAGAACGAAGGGCGTCGAAGCCCTGTAAAGCGGCGGTGCAACAATC 420
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGGATGCCA 480
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGGATGCCA 480
481 TTGCTGTGGAAGCTGCTGCGCATTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
481 TTGCTGTGGAAGCTGCTGCGCATTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
601 TGGTGGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGGCCATTTAAGTCTCTGCTCGG 660
601 TGGTGGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGGCCATTTAAGTCTCTGCTCGG 660
661 CGCGTCTGCGTCTGCTGGCTGGCATTAATACTCACTCGCAATCAAAATTCAGCCGATAG 720
661 CGCGTCTGCGTCTGCTGGCTGGCATTAATACTCACTCGCAATCAAAATTCAGCCGATAG 720
721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGTCTTTCACAAACCATGCAAAATGCTGA 780
721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGTCTTTCACAAACCATGCAAAATGCTGA 780
781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCATCAAAAGGATTTTC 960
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCATCAAAAGGATTTTC 960
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCATCAAAAGGATTTTC 960
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCATCAAAAGGATTTTC 960
961 GCCTGCTGGGCAACACCGAGTGGAGCCGCTTCTGCACTCTCTCAGGSCCAGGCGGTGA 1020
961 GCCTGCTGGGCAACACCGAGTGGAGCCGCTTCTGCACTCTCTCAGGSCCAGGCGGTGA 1020
1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATA 1080
1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATA 1080
1081 CGCAACCGCCTCTCCCGCGGTGGCGGATTCATTAATGAGCTGGGCACGACAGGTTT 1140
|||||

Db 1081 CGCAAAACCGCCCTCTCCCGCGCGTGGCCGATTCATTAATGAGCTGGCAGCAGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATTTCTATGTTTGACAGCTTATCATCGACTGCGACGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTTCTATGTTTGACAGCTTATCATCGACTGCGACGGTGCAACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCCATTCGGAAGCTGTGGTGTGTCAGGTGCGTAAATCACTGCGATAAATTCG 1320
Db 1261 TCAGGACGCCATTCGGAAGCTGTGGTGTGTCAGGTGCGTAAATCACTGCGATAAATTCG 1320
QY 1321 TGTGCTCAAGCGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTCGACAAATTAATCATCGGCTCGTAAATGCTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTCGACAAATTAATCATCGGCTCGTAAATGCTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTTAAACGGCGATTAAGGGCTATAA CCGTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAACGGCGATTAAGGGCTATAA CCGTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATTCGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATTCGGAAGAGAAATTCGCCAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGTGGCTGAATAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAATAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAATGAAATGCAAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAATGAAATGCAAGACACCGATTAC 2160
QY 2161 TCCATCGGAGAAAGCTGCTTTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGGAGAAAGCTGCTTTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220

```
QY 2221 GCATGGTCCAAATCGACACCCAGCAAGTGAATTAATGGTGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCCAGCAAGTGAATTAATGGTGTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACCTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACCTGATGAAGTCTG 2400
QY 2401 GAAGCGGTAAATAAAGACAAACCCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTAAATAAAGACAAACCCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGTTATGCGCCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATGCGCCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGAGATGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGAGATGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACACACACATAACATAACAAACACCTCGGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACACACACACATAACATAACAAACACCTCGGGATCGAGGAGGATTCAGAAATTC 2700

RESULT 10
AR031992 LOCUS AR031992 7475 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5866684.
ACCESSION AR031992
VERSION AR031992.1 GI:5946281
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 7475)
Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,F.Xavier.
TITLE Peptidyl inhibitors of viral proteases
JOURNAL Patent: US 5866684-A 1 02-FEB-1999;
FEATURES
source Location/Qualifiers
1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGTCAAAACCTTTCGCGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGTCAAAACCTTTCGCGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAAACAGTAAGCTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAAACAGTAAGCTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGCTGTGTAACACAGGCCAGCCAGTTCTTCGCAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGCTGTGTAACACAGGCCAGCCAGTTCTTCGCAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 240
```

```
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCCGCTGCAAAATTTGTCGGCGGATTAATCTCTCGCGCGATCAACTGGGTGGCGACGC 360
Db 301 ACGCCCGCTGCAAAATTTGTCGGCGGATTAATCTCTCGCGCGGATCAACTGGGTGGCGACGC 360
QY 361 TGTGTGTGTCGATGTAGAAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGTGTGTGTCGATGTAGAAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACCTAACTGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACCTAACTGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGATTAATTTCTCCATGAAGACGGTACGCGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGATTAATTTCTCCATGAAGACGGTACGCGACTGGCGTGGAGCATC 600
QY 601 TGGTCGCAATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAGTCTGTCTCGG 660
Db 601 TGGTCGCAATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAGTCTGTCTCGG 660
QY 661 CGGCTCTGCGCTGTCGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db 661 CGGCTCTGCGCTGTCGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
QY 721 CGGAAACGGGAAGCGCACTGGAAGTGCATGTCGGTTCCTCAACAAACCATCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGCGCACTGGAAGTGCATGTCGGTTCCTCAACAAACCATCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGCGCATTAACGAGTCCGGCTGCGGTTCGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCGCATTAACGAGTCCGGCTGCGGTTCGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACATACCGGAAGACAGCTCATGTTATATCCGCGCTTAACCAACATCAAAACAGGATTTTC 960
Db 901 ACATACCGGAAGACAGCTCATGTTATATCCGCGCTTAACCAACATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTGCTGCAACTCTCTCAGGGCGAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTGCTGCAACTCTCTCAGGGCGAGCGGTGA 1020
QY 1021 AGGCAATCAGTGTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGTGTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCGAGCTGGCACGACAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCGAGCTGGCACGACAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGAGTTAGTCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGAGTTAGTCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAGCCATCGGAAGCTGTTGATGCTGTGAGGCTGTAAATCACTGCATTAATTCG 1320
Db 1261 TCAGGCAGCCATCGGAAGCTGTTGATGCTGTGAGGCTGTAAATCACTGCATTAATTCG 1320
```


Qy	1321	TGTCCTCAAGCGGCACTCCCGCTTCGGATATATGTTTTTGGCGCGACATCATACGGTT	1380
Db	1321	TGTCCTCAAGCGGCACTCCCGCTTCGGATATATGTTTTTGGCGCGACATCATACGGTT	1380
Qy	1381	CTGGCAAAATATTTCTGAATAGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
Db	1381	CTGGCAAAATATTTCTGAATAGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
Qy	1441	ATTGTAGCGGATACAAATTTTACACAGGAAACAGCCAGTCGGTTAGTGTTTTTACGA	1500
Db	1441	ATTGTAGCGGATACAAATTTTACACAGGAAACAGCCAGTCGGTTAGTGTTTTTACGA	1500
Qy	1501	GCATTTCACCAACAGGACCATAGATTATGAAAACTGAAGAAAGTAAACTGGTATCTGG	1560
Db	1501	GCATTTCACCAACAGGACCATAGATTATGAAAACTGAAGAAAGTAAACTGGTATCTGG	1560
Qy	1561	ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT	1620
Db	1561	ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT	1620
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAGAGAAATTCGCCAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAGAGAAATTCGCCAGGTT	1680
Qy	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACGGTTTGGTGGCTAC	1740
Db	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACGGTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCTCTGGCTGAATCAACCCGGACAAAGGTTCCAGGACAAAGCTGAT	1800
Db	1741	GCTCAATCTGGCTCTGGCTGAATCAACCCGGACAAAGGTTCCAGGACAAAGCTGAT	1800
Qy	1801	CGTTTACCTGGATGCGGTAGCTTACAAACGCAAGCTGATTCCTTACCCGATCCTGTT	1860
Db	1801	CGTTTACCTGGATGCGGTAGCTTACAAACGCAAGCTGATTCCTTACCCGATCCTGTT	1860
Qy	1861	GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
Db	1861	GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
Qy	1921	GAGATCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTTAGAGCGCGTGAATTCAC	1980
Db	1921	GAGATCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTTAGAGCGCGTGAATTCAC	1980
Qy	1981	CTGCAAGAACCGTACTTCACTGGCGGCTGATTCGCTGACGGGGTTATCGCTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGGCGGCTGATTCGCTGACGGGGTTATCGCTTCAAG	2040
Qy	2041	TATGAAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG	2100
Qy	2101	GGTCTGACCTTCTGGTTGACCTGATTAATAACCAACACATCAATCAACGCGCGATAC	2160
Db	2101	GGTCTGACCTTCTGGTTGACCTGATTAATAACCAACACATCAATCAACGCGCGATAC	2160
Qy	2161	TCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATCAACATCAACGCGCGGTTG	2220
Db	2161	TCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATCAACATCAACGCGCGGTTG	2220
Qy	2221	GCATGGTCCAAATCGACACCAAGAGTGAATTTATGGTGAACGGTACTGCGGACCTTC	2280
Db	2221	GCATGGTCCAAATCGACACCAAGAGTGAATTTATGGTGAACGGTACTGCGGACCTTC	2280
Qy	2281	AAGGTCAACCATCAACCGTTGTTGGCTGCTGAGCGCAGGATTAATAACGCGCGCAGT	2340
Db	2281	AAGGTCAACCATCAACCGTTGTTGGCTGCTGAGCGCAGGATTAATAACGCGCGCAGT	2340
Qy	2341	CGAAACAAAGAGCTGCGAAACAGTTCCTCGAAACTATCTGACTGATGAAGGCTCG	2400
Db	2341	CGAAACAAAGAGCTGCGAAACAGTTCCTCGAAACTATCTGACTGATGAAGGCTCG	2400
Qy	2401	GAAGCGTTAATAAGACAAACCCGCTGGTCCGTAGCGCTGAAGTCTTACGAGGAAG	2460
Db	2401	GAAGCGTTAATAAGACAAACCCGCTGGTCCGTAGCGCTGAAGTCTTACGAGGAAG	2460
Qy	2461	TTGGCGAAAGATCCACGTATTGGCCGACCATGGAACCGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTATTGGCCGACCATGGAACCGCCAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGTCGGTCTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGTCGGTCTGATCAACGCC	2580
Qy	2581	GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAGACGCGACACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAGACGCGACACTAAATTCGAGCTCG	2640
Qy	2641	AACAACAAACAATAACAATAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC	2700
Db	2641	AACAACAAACAATAACAATAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC	2700
RESULT 11			
AR207294			
LOCUS	AR207294	7475 bp	DNA linear PAT 20-JUN-2002
DEFINITION	Sequence 1 from patent US 6372883.		
ACCESSION	AR207294		
VERSION	AR207294.1	GI:21506162	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 7475)		
AUTHORS	Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen., Kay,P.Brittain., Raynham,T.Michael. and Wilson,F.Xavier.		
TITLE	Antiviral medicaments		
JOURNAL	Patent: US 6372883-A 1 16-APR-2002;		
FEATURES	Location/Qualifiers		
source	1..7475		
ORIGIN	/organism="unknown"		
	/mol_type="unassigned DNA"		
Query Match	81.8%; Score 2700; DB 6; Length 7475;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 2700; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATCGCATGATAGCGCCCGGAAGAGA	60
Db	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATCGCATGATAGCGCCCGGAAGAGA	60
Qy	61	GTCAATTTCAGGGTGGTGAATGTGAAACCAAGTAACGTTATACGATGTCGACAGTATGCG	120
Db	61	GTCAATTTCAGGGTGGTGAATGTGAAACCAAGTAACGTTATACGATGTCGACAGTATGCG	120
Qy	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTTCGGA	180
Db	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTTCGGA	180
Qy	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATTCATCCCAACCGCGTGGCAC	240
Db	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATTCATCCCAACCGCGTGGCAC	240
Qy	241	AACAACTCGCGGCAACAGCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTCGCGGCAACAGCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Qy	301	ACGCGCGTTCGCAATTCGCGCGGATTAATCTCGCGCGATCAATCGGGTGCAGCG	360
Db	301	ACGCGCGTTCGCAATTCGCGCGGATTAATCTCGCGCGATCAATCGGGTGCAGCG	360
Qy	361	TGGTGGTTCGATGCTAGAACGAGCGCGTTCGAGCTGTAAAGCGGGTGCACATC	420
Db	361	TGGTGGTTCGATGCTAGAACGAGCGCGTTCGAGCTGTAAAGCGGGTGCACATC	420

QY 421 TTCTCGCCAAACGGCTCAGTGGGCTGATCAITAACTATCCGCTGATGACCAAGATGCCA 480
DB |||||
QY 421 TTCTCGCCAAACGGCTCAGTGGGCTGATCAITAACTATCCGCTGATGACCAAGATGCCA 480
DB |||||
QY 481 TTGCTGTGGAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 540
DB |||||
QY 481 TTGCTGTGGAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 540
DB |||||
QY 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGATCGGCGCTGGAGCATC 600
DB |||||
QY 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGATCGGCGCTGGAGCATC 600
DB |||||
QY 601 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGGCCAATTAAGTTCTGTCTCGG 660
DB |||||
QY 601 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGGCCAATTAAGTTCTGTCTCGG 660
DB |||||
QY 661 CGCGTCTCGGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB |||||
QY 661 CGCGTCTCGGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB |||||
QY 721 CGGAACGGGAAGGACATGGAGTGCATGTCGGGTTTTCAACAACCAATGCAAAATGCTGA 780
DB |||||
QY 721 CGGAACGGGAAGGACATGGAGTGCATGTCGGGTTTTCAACAACCAATGCAAAATGCTGA 780
DB |||||
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGCTGTTGCCAACGATCGATGCGCTGGCGGCAA 840
DB |||||
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGCTGTTGCCAACGATCGATGCGCTGGCGGCAA 840
DB |||||
QY 841 TGGCGGCATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
DB |||||
QY 841 TGGCGGCATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
DB |||||
QY 901 ACGATACCGAAGACGATCATGTTATATCCCGCGGTTAAACCAACCATCAAAACGAGATTTTC 960
DB |||||
QY 901 ACGATACCGAAGACGATCATGTTATATCCCGCGGTTAAACCAACCATCAAAACGAGATTTTC 960
DB |||||
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB |||||
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB |||||
QY 1021 AGGGCAATCAGCTGTGCCCCCTCTCACTGGTGAAGAAAGAAACCAACCTGCGGCCAATA 1080
DB |||||
QY 1021 AGGGCAATCAGCTGTGCCCCCTCTCACTGGTGAAGAAAGAAACCAACCTGCGGCCAATA 1080
DB |||||
QY 1081 CGCAAAACGGCTCTCCCCCGCGGTTGGCCGATTCATTAAATGCACTGGCAGCAGAGTTT 1140
DB |||||
QY 1081 CGCAAAACGGCTCTCCCCCGCGGTTGGCCGATTCATTAAATGCACTGGCAGCAGAGTTT 1140
DB |||||
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
DB |||||
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGATGCAAGTGAAGTGAAGTGAAGTGAAGTGA 1260
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGATGCAAGTGAAGTGAAGTGAAGTGAAGTGA 1260
DB |||||
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGTAAATCACTGCAATAATTCG 1320
DB |||||
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGTAAATCACTGCAATAATTCG 1320
DB |||||
QY 1321 TGTGCTCAAGCGCATCTCCGGTCTCGGATTAATGTTTTTGGCGCGGACATCATCAACCGGTT 1380
DB |||||
QY 1321 TGTGCTCAAGCGCATCTCCGGTCTCGGATTAATGTTTTTGGCGCGGACATCATCAACCGGTT 1380
DB |||||
QY 1381 CTGGCAAAATTTCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
DB |||||
QY 1381 CTGGCAAAATTTCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
DB |||||
QY 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTACGA 1500
DB |||||
QY 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTACGA 1500
DB |||||

QY 1501 GCATTCACCAACAGGACCATAGATTATGAAAACCTGAAGAGTAAATCGGTAAATCTGG 1560
DB |||||
QY 1501 GCATTCACCAACAGGACCATAGATTATGAAAACCTGAAGAGTAAATCGGTAAATCTGG 1560
DB |||||
QY 1561 ATTAACCGCGGATAAAGGCTATAACGGTCTCGCTGAAAGTCCGGTAAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1561 ATTAACCGCGGATAAAGGCTATAACGGTCTCGCTGAAAGTCCGGTAAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1621 ACCGGAATTAAGGTCACCGTTGAGCATCCCGATAAACTGGAAGAGAAATTTCCCAACGGTT 1680
DB |||||
QY 1621 ACCGGAATTAAGGTCACCGTTGAGCATCCCGATAAACTGGAAGAGAAATTTCCCAACGGTT 1680
DB |||||
QY 1681 GCGGCAATCTGGCGATGCGCTGACATTAATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740
DB |||||
QY 1681 GCGGCAATCTGGCGATGCGCTGACATTAATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740
DB |||||
QY 1741 GCTCAATCTGGCGCTGTTGGCTGAAATCAACCCCGCAAAAGGTTCCAGGACAAAGCTGTAT 1800
DB |||||
QY 1741 GCTCAATCTGGCGCTGTTGGCTGAAATCAACCCCGCAAAAGGTTCCAGGACAAAGCTGTAT 1800
DB |||||
QY 1801 CCGTTTACCTGGGATGCGCTGACGTTAAACGGAAGCTGATTGCTTACCCGATCGCTGTT 1860
DB |||||
QY 1801 CCGTTTACCTGGGATGCGCTGACGTTAAACGGAAGCTGATTGCTTACCCGATCGCTGTT 1860
DB |||||
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
DB |||||
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
DB |||||
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTGCTGCTGACGGGGGTTATGGTTCAAG 2040
DB |||||
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTGCTGCTGACGGGGGTTATGGTTCAAG 2040
DB |||||
QY 2041 TATGAAACCGGCAAGTACGACATTAAGAGCTGGGCGTGGATTAACGCTGGCGGAAACG 2100
DB |||||
QY 2041 TATGAAACCGGCAAGTACGACATTAAGAGCTGGGCGTGGATTAACGCTGGCGGAAACG 2100
DB |||||
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGAATGCAGACACCGGATTAC 2160
DB |||||
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGAATGCAGACACCGGATTAC 2160
DB |||||
QY 2161 TCCATCGCAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGGTG 2220
DB |||||
QY 2161 TCCATCGCAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGGTG 2220
DB |||||
QY 2221 GCATGFTCCAACTCGACACCAAGTGAATTTATGTTGTAACGGTACTTCCGACCTTC 2280
DB |||||
QY 2221 GCATGFTCCAACTCGACACCAAGTGAATTTATGTTGTAACGGTACTTCCGACCTTC 2280
DB |||||
QY 2281 AAGGTCACCATCCAAACCGTTGTTGGGCTGCTGAGCGCAGGTATTAAACGGCCGAGT 2340
DB |||||
QY 2281 AAGGTCACCATCCAAACCGTTGTTGGGCTGCTGAGCGCAGGTATTAAACGGCCGAGT 2340
DB |||||
QY 2341 CCGAAACAAAGAGCTGCGCAAAAGGTTCTCGAAAACTATCTGCTGATGAAGGTCG 2400
DB |||||
QY 2341 CCGAAACAAAGAGCTGCGCAAAAGGTTCTCGAAAACTATCTGCTGATGAAGGTCG 2400
DB |||||
QY 2401 GAAGCGGTTTAAAGAGCAAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
DB |||||
QY 2401 GAAGCGGTTTAAAGAGCAAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
DB |||||
QY 2461 TTGGGAAAGATCCAGTATTGCGGCAACCATGGAAAAACCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2461 TTGGGAAAGATCCAGTATTGCGGCAACCATGGAAAAACCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2521 CCGAAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
DB |||||
QY 2521 CCGAAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
DB |||||
QY 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640

Db 2581 GCGAGGGTCGTGACAGTGTGATGAGCCCTGAAAGACGCCGACAGTAATTCGAGCTCG 2640

Qy 2641 AACAAACAACAATAACAATAACAACACCTCGGGATGCGAGGAGGATTTGAGAATTC 2700

Db 2641 AACAAACAACAATAACAATAACAACACCTCGGGATGCGAGGAGGATTTGAGAATTC 2700

RESULT 12

AX377532

LOCUS AX377532 9191 bp DNA linear PAT 18-MAR-2002

DEFINITION Sequence 9 from Patent WO0212553.

ACCESSION AX377532

VERSION AX377532.1 GI:19573718

KEYWORDS .

SOURCE synthetic construct

ORGANISM synthetic construct

other sequences; artificial sequences.

REFERENCE 1

AUTHORS Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdoerf,H. and Muth,J.

TITLE Method for detecting mutations in nucleotide sequences

JOURNAL Patent: WO 0212553-A 9 14-FEB-2002;

Nanogen Recognomics GmbH (DE)

FEATURES

source

1. .9191

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="pMALc2x-mutS Plasmid"

ORIGIN

Query Match 81.8%; Score 2698; DB 6; Length 9191;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CCGACACCATCAATGGTGCACAAACCTTTCGGGGTATGCGCATGATAGCCCGGAAGAGA 60

Db 1 CCGACACCATCAATGGTGCACAAACCTTTCGGGGTATGCGCATGATAGCCCGGAAGAGA 60

Qy 61 GTCAATTCAAGGTGGTGAATGTGAAACAGTAACGTTATACCATGTGCGAGAGTATGCCG 120

Db 61 GTCAATTCAAGGTGGTGAATGTGAAACAGTAACGTTATACCATGTGCGAGAGTATGCCG 120

Qy 121 GTGTCTCTTATCAGACCGGTTTCCTCCGGTGGTGAACCCAGGCCACGCTTTCCTGCGAAAA 180

Db 121 GTGTCTCTTATCAGACCGGTTTCCTCCGGTGGTGAACCCAGGCCACGCTTTCCTGCGAAAA 180

Qy 181 CCGGGGAAAAAGTGGNAGCGCGATGCGGAGCTGAAATTACATTCGCAACCGCTGGGCAC 240

Db 181 CCGGGGAAAAAGTGGNAGCGCGATGCGGAGCTGAAATTACATTCGCAACCGCTGGGCAC 240

Qy 241 AACAACTGGCGGCAACAGTGTGCTGATTGGGGTTGCCACCTCCAGTCTGGCCCTGC 300

Db 241 AACAACTGGCGGCAACAGTGTGCTGATTGGGGTTGCCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGCGCGTCGCAAAATGTTCGGGGGATTAATATCTCGCCCGATCAACTGGGTGCCAGCG 360

Db 301 ACGCGCGTCGCAAAATGTTCGGGGGATTAATATCTCGCCCGATCAACTGGGTGCCAGCG 360

Qy 361 TGGTGTGTGATGTTAGAACGAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAATTC 420

Db 361 TGGTGTGTGATGTTAGAACGAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAATTC 420

Qy 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCTATCCGTGGATGACCAAGGTGCCA 480

Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCTATCCGTGGATGACCAAGGTGCCA 480

Qy 481 TTGCTGTGGAAGTGCCTGCATTAATGTTCCGGCGTTATTTCTGATGTCTCTGACCGAGA 540

Db 481 TTGCTGTGGAAGTGCCTGCATTAATGTTCCGGCGTTATTTCTGATGTCTCTGACCGAGA 540

Qy 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGACCTGGCGGTGGAGCATC 600

Db 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600

Qy 601 TGGTCGCATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGGCCCAATTAAGATTCTGTCTCGG 660

Db 601 TGGTCGCATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGGCCCAATTAAGATTCTGTCTCGG 660

Qy 661 CGCGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Db 661 CGCGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCACTGCAAAATGCTGA 780

Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCACTGCAAAATGCTGA 780

Qy 781 ATGAGGGCATCGTTCCCACTCGCATGTGTTGCAACGATCAGATGCGCTGGGCGCAA 840

Db 781 ATGAGGGCATCGTTCCCACTCGCATGTGTTGCAACGATCAGATGCGCTGGGCGCAA 840

Qy 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900

Db 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900

Qy 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACCATCAAAACAGGATTTTC 960

Db 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACCATCAAAACAGGATTTTC 960

Qy 961 GCCTGCTGGGCAAAACCAAGCGTGGACCGCTTCTCTCAGGCGCCAGCGGTGA 1020

Db 961 GCCTGCTGGGCAAAACCAAGCGTGGACCGCTTCTCTCAGGCGCCAGCGGTGA 1020

Qy 1021 AGGGCAATCAGCTGTGTCCTCTCATCTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080

Db 1021 AGGGCAATCAGCTGTGTCCTCTCATCTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140

Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140

Qy 1141 CCCGACTCGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200

Db 1141 CCCGACTCGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200

Qy 1201 GCACAAATTCATGTTTGCACAGCTTATCATCGATGCAACGCGTGCAACCAATGCTTCGGCG 1260

Db 1201 GCACAAATTCATGTTTGCACAGCTTATCATCGATGCAACGCGTGCAACCAATGCTTCGGCG 1260

Qy 1261 TCAGGACGCGATCGGAAGCTGTGGTATGCTGTGAGGTGCTGATCACTGCATTAATTCG 1320

Db 1261 TCAGGACGCGATCGGAAGCTGTGGTATGCTGTGAGGTGCTGATCACTGCATTAATTCG 1320

Qy 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCATACCGGTT 1380

Db 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCATACCGGTT 1380

Qy 1381 CTGGCAAAATATCTGAAATAGCTGTTGCAATTAATCATCGGCTCGTATTAATGTGTGA 1440

Db 1381 CTGGCAAAATATCTGAAATAGCTGTTGCAATTAATCATCGGCTCGTATTAATGTGTGA 1440

Qy 1441 ATTGTAGCGGNTAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA 1500

Db 1441 ATTGTAGCGGNTAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA 1500

Qy 1501 GCACCTTACCAACAGGACCATAGATTAATGAAGAGGTAAGTAACTGGTAACTCTGG 1560

Db 1501 GCACCTTACCAACAGGACCATAGATTAATGAAGAGGTAAGTAACTGGTAACTCTGG 1560

Qy 1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620

Db 1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGNTAACTGGAAGAGAAATTCGCCAGGTT 1680

Db	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGSAAAGAAATATCCCAACAGTTT	1681
Qy	1691	CGCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCCTGTTGGCTGAATATCAACCCCGGACAAAGCGTTTCCAGGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCCTGTTGGCTGAATATCAACCCCGGACAAAGCGTTTCCAGGACAAGCTGTAT	1800
Qy	1801	CCGTTTACCTCGGATGCCGTACGTTTAAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTCGGATGCCGTACGTTTAAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGTCGCCAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGTCGCCAACCCGCCAAAACCTGGGAA	1920
Qy	1921	GAGATCCGGCGCTGGATAAGAACTGAAAGCGAAAGGTAAGACGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGGCGCTGGATAAGAACTGAAAGCGAAAGGTAAGACGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTCACTGGCCGCTGATTGCTGCTGACCGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGGCCGCTGATTGCTGCTGACCGGGGTTATGCGTTCAAG	2040
Qy	2041	TATCAAAACGGCAAGTACGACATTAAAGACGTTGGCGCTGGATACGCTGCGCGGAAAGCG	2100
Db	2041	TATCAAAACGGCAAGTACGACATTAAAGACGTTGGCGCTGGATACGCTGCGCGGAAAGCG	2100
Qy	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAAACATGAATGAGACACCGGATTAC	2160
Db	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAAACATGAATGAGACACCGGATTAC	2160
Qy	2161	TCCATCGAGAGCTGCTTTTATAAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG	2220
Db	2161	TCCATCGAGAGCTGCTTTTATAAAGGCGGAAACAGCGATGACCATCAACGCGCCGCTGG	2220
Qy	2221	GCATGTTCCAACTCGACACAGCAAAAGTGAATTTATGGTCTTAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGTTCCAACTCGACACAGCAAAAGTGAATTTATGGTCTTAACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGTCACCATCTCAAAACCGTTTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT	2340
Db	2281	AAGGTCACCATCTCAAAACCGTTTTCGTTGGCGTCTGAGCGCAGGTATTAACGCGCCAGT	2340
Qy	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGTAAGTGAAGTCTG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGTAAGTGAAGTCTG	2400
Qy	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGGAAGATCCAAGTATTTGGCGCCACATCGAAAAACGCCAGAAAAAGTGAATCATG	2520
Db	2461	TTGGCGGAAGATCCAAGTATTTGGCGCCACATCGAAAAACGCCAGAAAAAGTGAATCATG	2520
Qy	2521	CCGAAATCCCGCAGATGTCGGCTTTCTGGTATGTCGCTGGCTACTCGGCTGATCAACGCC	2580
Db	2521	CCGAAATCCCGCAGATGTCGGCTTTCTGGTATGTCGCTGGCTACTCGGCTGATCAACGCC	2580
Qy	2581	GCAGCGGCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCAGCGGCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAAACAAACAATAACCAATAACCAACCTCGGATCGAGGGAGGATTTTCAGAAATTC	2700
Db	2641	AACAAACAAACAATAACCAATAACCAACCTCGGATCGAGGGAGGATTTTCAGAAATTC	2700
Qy	2701	GGATCC	2706
Db	2701	GGATCC	2706

gene	/note="multiple cloning site"			
promoter	3426..4356			
	/gene="bla"			
	3426..3454			
	/gene="bla"			
CDS	3496..4356			
	/gene="bla"			
	/codon_start=1			
	/transl_table=11			
ORIGIN	/product="beta-lactamase"			
	/protein_id="AAC83814.1"			
	/db_xref="GI:3983123"			
	/translation="MSIOHPRVALIPFAFLCPVFAHPETLVKVKDAEDOLGARVGY			
	IELDLSKILLESFRRPERSPMWTFKVLICGAVLSRYDAGQQLGRFLHYSDNDLVE			
	YSPVTEKHLTDGMTVRELCSAAITWSDNTANLLLTITGGPKELTAFIHNMGDHDVRL			
	DRWPELNEAIPNDRDTPMPVAMATILRLKLLTGELLTLASRQQLIDWMEADKVGPL			
	LRSLAPAGWFIADKSGAGERSGRIIAALGPDGKPSRIWVITYTGSQATMDERNRQIA			
	EIGASLIKHW"			
Query Match	81.1%;	Score 2676.2;	DB 12;	Length 6724;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 2678;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	CGGACACCATCGAATGGTGCAAAACCTTTCGGGTATGGCATAGTAGCGCCGGAAGAGA	60	
DB	1	CGGACACCATCGAATGGTGCAAAACCTTTCGGGTATGGCATAGTAGCGCCGGAAGAGA	60	
QY	61	GTCAATTACAGGCTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGACAGATGATGCCG	120	
DB	61	GTCAATTACAGGCTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGACAGATGATGCCG	120	
QY	121	GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGGCCAGCCACGTTTCTGCGAAAA	180	
DB	121	GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGGCCAGCCACGTTTCTGCGAAAA	180	
QY	181	CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240	
DB	181	CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240	
QY	241	AACAACTGCGCGGCAAAACAGTCTGTGCTGATTGGCGTTGGCCACCTCCAGTCTGGCCCTGC	300	
DB	241	AACAACTGCGCGGCAAAACAGTCTGTGCTGATTGGCGTTGGCCACCTCCAGTCTGGCCCTGC	300	
QY	301	ACGCGCCGTTCGCAAAATGTCGCGCGGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCG	360	
DB	301	ACGCGCCGTTCGCAAAATGTCGCGCGGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCG	360	
QY	361	TGGTGGTGCAGTGTAGAAAGAAAGCGGCTCGAAGCCCTGTAAAGCGGCGGTGCACAATC	420	
DB	361	TGGTGGTGCAGTGTAGAAAGAAAGCGGCTCGAAGCCCTGTAAAGCGGCGGTGCACAATC	420	
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACTATCCGCTGGATGACAGGATGCCA	480	
DB	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACTATCCGCTGGATGACAGGATGCCA	480	
QY	481	TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA	540	
DB	481	TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA	540	
QY	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC	600	
DB	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC	600	
QY	601	TGGTGCATTTGGTCAACAGAAATCGCGCTGTAGCGGGCCCATTAAGTCTGCTCGG	660	
DB	601	TGGTGCATTTGGTCAACAGAAATCGCGCTGTAGCGGGCCCATTAAGTCTGCTCGG	660	
QY	661	CGCGTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
DB	661	CGCGTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
QY	721	CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAATGCTGA	780	

Db 1801 CCGTTTACCTGGATGCCGTAAGTTACAAACGCAAGCTGATTCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGTGTATTTATACAAAGATCTGTCGCAACCCGCCAARACCTGGGA 1920
Db 1861 GAAGCGTTATCGTGTATTTATACAAAGATCTGTCGCAACCCGCCAARACCTGGGA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGGTAAAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGGTAAAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCTACTTACCTGGCGCTGATTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCTACTTACCTGGCGCTGATTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTCGGCTGGATAACGCTGGCGCAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTCGGCTGGATAACGCTGGCGCAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACACGAGCAAGTGAATTAATGTTAAACGCTGACCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGAGCAAGTGAATTAATGTTAAACGCTGACCGACCTTC 2280
QY 2281 AAGGTCACACCATCAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGGCCCGCAGT 2340
Db 2281 AAGGTCACACCATCAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGGCCCGCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAACATCTCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAACATCTCTGACTGATGAAGGCTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGTTATTCGCCACCATGAAACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATTCGCCACCATGAAACGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTATGCGTGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTATGCGTGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACACAAACAATAACAATAACAACACCTCGGGATCGA 2681
Db 2641 AACACACAAACAATAACAATAACAACACCTCGGGATCGA 2681

RESULT 14
AX378208
LOCUS AX378208 6806 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 194 from Patent WO0206306.
ACCESSION AX378208
VERSION AX378208.1 GI:19574061
KEYWORDS
SOURCE
ORGANISM
unidentified
unidentified
unclassified.
REFERENCE
1
Yan, R., Tomasselli, A. G., Gurney, M. E., Emmons, T. L., Bienkowski, M. J.
and Heinriksen, R. L.
Substrates and assays for (b)-secretase activity

JOURNAL Patent: WO 0206306-A 194 24-JAN-2002;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1..6806
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="fusion protein comprising a maltose binding protein
with 125 amino"

ORIGIN
Query Match 79.9%; Score 2636.8; DB 6; Length 6806;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2638; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCGACACATCGAATGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACATCGAATGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGA 60
QY 61 GTCAATTTCAGGTGTGAATGTGAACACAGTAAGTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTTCAGGTGTGAATGTGAACACAGTAAGTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGCTGTGAACACGAGCCAGCCAGCTTTCTGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGCTTTCCCGCTGTGAACACGAGCCAGCCAGCTTTCTGCGAAA 180
QY 181 CGCGGAAAAGTGAAGCGCGGATGCGGAGCTGAATTAATTCACATTCACACCGCGTGGCAC 240
Db 181 CGCGGAAAAGTGAAGCGCGGATGCGGAGCTGAATTAATTCACATTCACACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTCGTTGCTGATTGCGCGTTCGCACTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCGTTGCTGATTGCGCGTTCGCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATGTGTCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATGTGTCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGTGCTGTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGGCGTGCACAATC 420
Db 361 TGTGCTGTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGGCGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCATTAATCATTCGCTGGATGACCGAGTGGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCATTAATCATTCGCTGGATGACCGAGTGGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTTCGCGCTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTTCGCGCTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAGACGCTAGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAGACGCTAGCGACTGGGCGTGGAGCATC 600
QY 601 TGTGTCGATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db 601 TGTGTCGATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
QY 661 CGGCTCTGCTGTGGCTGGCTGGCATAAATATCTCATCTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGCTGTGGCTGGCTGGCATAAATATCTCATCTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACCGGAAGCGCACTGGAGTGGCATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 CGGAACCGGAAGCGCACTGGAGTGGCATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 781 ATGAGGCAATCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 ATGAGGCAATCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 TCGCGCGCAATTCAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 900
Db 841 TCGCGCGCAATTCAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 900

Db 841 TCGCGCCATTACCGAGTCCGGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATAGG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCAATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACCGTGGACCGTTCGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACCGTGGACCGTTCGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGGCCGCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGGCCGCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGACGTGGCAGCAGAGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGACGTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGACGCAACGCAATTAATGTAGTTCAGTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGACGCAACGCAATTAATGTAGTTCAGTCACTCATTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATCACTGCAAGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATCACTGCAAGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGCGAGGTGTAATCACTGATTAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGCGAGGTGTAATCACTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCGCTTCTGATTAATGTTTTTGGCGCGACATCATACAGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCGCTTCTGATTAATGTTTTTGGCGCGACATCATACAGGTT 1380
Qy 1381 CTGGCAATATCTGAAATGAGCTGTGACATTAATCATCGGCTGATATGTTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTGACATTAATCATCGGCTGATATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTACACAGGAAACAGCCAGTCCGTTAGGTGTTTTCAGGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTACACAGGAAACAGCCAGTCCGTTAGGTGTTTTCAGGA 1500
Qy 1501 GCACCTCACCAAGGACCATAGATTATGAAGAACTGAAGAGGTAAACTGGTAAATCTGG 1560
Db 1501 GCACCTCACCAAGGACCATAGATTATGAAGAACTGAAGAGGTAAACTGGTAAATCTGG 1560
Qy 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGCTTACGATCCGGAATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGCTTACGATCCGGAATAAATCGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGACACGACCGCTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGACACGACCGCTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCCGGCAAGCGTTCCAGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCCGGCAAGCGTTCCAGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATCAAGATCTGCTCCGAAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATCAAGATCTGCTCCGAAACCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATGAAGAACTGAAGCGAAAGGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATGAAGAACTGAAGCGAAAGGTGAAGCGCGCTGATGTTCAAC 1980

Qy 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATGCTGTCGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATGCTGTCGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAGACGCAAGTACGACATTAAGACCTGGGCGTGGATTAAGCTGGCGGAAAGCG 2100
Db 2041 TATGAAGACGCAAGTACGACATTAAGACCTGGGCGTGGATTAAGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAGAAACAAACATGATGACAGACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAGAAACAAACATGATGACAGACCGATTAC 2160
Qy 2161 TCCATCGAAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCGAAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACGCAAGCAAGTGAATTAATGTTGTAACGCTACTGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACGCAAGCAAGTGAATTAATGTTGTAACGCTACTGCGACCTTC 2280
Qy 2281 AAGGCTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGCTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAACGGTTAATAAGACAAACCGCTGGTCCGCTAGCGTGAAGTCTTACGAGGAAG 2460
Db 2401 GAACGGTTAATAAGACAAACCGCTGGTCCGCTAGCGTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCAAGAGTCCACGCTATTGCGCCCAACCATGGAAGAAACGCCAGAAAGTGAAATCATG 2520
Db 2461 TTGGCAAGAGTCCACGCTATTGCGCCCAACCATGGAAGAAACGCCAGAAAGTGAAATCATG 2520
Qy 2521 CCGAACAATCCCGCAGATGTCGCTTTTCTGGTATGCCGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACAATCCCGCAGATGTCGCTTTTCTGGTATGCCGCTGCTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTCAGACTGCTGATGAAGCTTGAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTCAGACTGCTGATGAAGCTTGAAGACGCGCAGACTTAATTCGAGCTCG 2640

RESULT 15
AF031088 6706 bp DNA circular SYN 07-MAY-2001
LOCUS Shuttle vector pMAL-pIII, complete sequence.
DEFINITION AF031088
ACCESSION AF031088
VERSION AF031088.1 GI:2623821
KEYWORDS Shuttle vector pMAL-pIII
SOURCE Shuttle vector pMAL-pIII
ORGANISM Shuttle vector pMAL-pIII
OTHER SEQUENCES; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6706)
AUTHORS Zwick,M.B., Bonnycastle, L.L., Noren, K.A., Venturini, S., Leong, E.,
Barbas, C.F. III, Noren, C.J. and Scott, J.K.
TITLE The maltose-binding protein as a scaffold for monovalent display of
peptides derived from phage libraries
JOURNAL Anal. Biochem. 264 (1), 87-97 (1998)
MEDLINE 9902881
PUBMED 9784192
REFERENCE 2 (bases 1 to 6706)
AUTHORS Noren, K.A. and Noren, C.J.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road,
Beverly, MA 01915, USA
FEATURES Location/Qualifiers
source 1..6706
/organism="Shuttle vector pMAL-pIII"
/mol_type="genomic DNA"

```

/db_xref="taxon:69327"
/Note="derived from commercially available vector pMal-p2
(New England Biolabs)"
15..1163
/Name="lacI"
/Note="E. coli gene"
15..46
/Name="lacI"
/Name="lacI"
81..1163
/Name="lacI"
/Note="transcriptional repressor"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAB86558.1"
/db_xref="GI:2623822"
/translation="MKPVTLYDVAEYAGVSYQTVSRVUNQASHVSAKTRKVEAMAB
LNYIPNRVAQLAGKQSLIGVATSSLAHAFSQIVAAIKRADQLGASVVSVMS
GVEACKAAHNLQARVGLIINPLDQDAIEAECTNVPALFLVDSDQTPINSEI
FSHEDGTGLVHLVALGQIALLAGLSVSARLRAGLWHKYLTRNQIPIAEREG
DWSAMSGFQMTQMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADI SVVGYDDT
EDSCYIPPLATTIKQDFRLGQTSVDRLLQLSQAVKGNQLLPVSLVKRKTLLPNT
QTASPRALADSLMLQARQVSRLESQ"
1406..1433
/Note="PtaC promoter"
1528..2976
/Name="malE-lacZa fusion"
/Note="E. coli gene fusion"
1528..2976
/Name="malE-lacZa fusion"
/Note="MBP-lacZa; MBP modified to include phage M13
secretion signal sequence"
/codon_start=1
/transl_table=11
/product="maltose binding protein-lacZ alpha peptide
fusion protein precursor"
/protein_id="AAB86559.1"
/db_xref="GI:2623823"
/translation="MKKLLFAIPLVVPFYSHSDMKIEGKLVIIWINDKGYNGLAHV
GKFEQDTGIVTVVEHPDKLEKPEQVAATGDDPDII FWAHDFPGVQAQGLLAEITP
DKAFQDKLPTVDVYNGKLIAYPIAVEALSILYINKOLLPNPKTWBEI PALDKEL
KAGKSAIMFNLPQYPFTPLIIAADGGVAFKYGKVDI KDVGVDNAGAKAGLFLVD
LKNKHNADTDYSIAEAPNKGETANTINGPWNANIDISKVNYGVTVLPFTKGQDA
KPPFVLSAGINASPENKELAKFENLYLIDEGLEAVNKDKPLGVALKLSBEEELAK
DPRIAATHENAKQIEIPNFIQMSAFWIAVTVINAASGRQTVDEALDKDAQTNSSN
NNNNNNNLGIEIRISFEFGSSRVLDQASLALAVLQRRDWNENPGVTQLNRLAAHPFF
ASWRNSEARTDRPSQQLRSUNGWQLGCFGG"
1528..1581
/Name="malE-lacZa fusion"
/Note="M13 secretion signal sequence"
1582..2973
/Name="malE-lacZa fusion"
/product="maltose binding protein-lacZ alpha peptide
fusion protein"
3478..4338
/Name="bla"
/Note="R-plasmid ampicillin resistance gene"
3478..4338
/Name="bla"
/Note="penicillinase"
/codon_start=1
/transl_table=11
/product="RTEW-1 beta-lactamase"
/protein_id="AAB86560.1"
/db_xref="GI:2623824"
/translation="MSIQHPRVLAIPFFAAFCPLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILDESFRPEERSPMSTFKVLICGLAVLSKRVDAQEQQLGRIIHYSDNLAVE
YSPVTEKHLTDGMTVERELCSAAITMSDNTAANLLTTIGGPKELTAFLHMGDHYTRL
DRWEPELNEALPNDERDTPPVAMATTLRKLLTGELLTLASRQQLIDFWWDKVAQPL
LRALPAGWFIADKSGNGERSGRIIAALGPDGKPSRIVVITYTTSQATWDERNRQIA
EIGASLIKHW"
4381..4893
/standard_name="M13 origin (+ -)"

```

```

ORIGIN
Query Match      79.7%; Score 2631; DB 12; Length 6706;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 60; Gaps 1;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGGCCCGGAAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGGCCCGGAAAGAGA 60
QY 61 GTCAAATTCAAGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTGCGAGAGTATCCG 120
DB 61 GTCAAATTCAAGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTGCGAGAGTATCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCCAGCGTTCCTGCGAAAA 180
DB 121 GTGTCTCTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCCAGCGTTCCTGCGAAAA 180
QY 181 CCGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
DB 181 CCGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCGTCGCAAAATGTCGCGCGGATTAATTCCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGGCGCGTCGCAAAATGTCGCGCGGATTAATTCCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGCGATGGTAGAACGAAGCGGCTGCAAGCCCTGTAAGCGCGGTGCACAATC 420
DB 361 TGGTGGTGTGCGATGGTAGAACGAAGCGGCTGCAAGCCCTGTAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTCCTCGCTGGATGACCAAGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTCCTCGCTGGATGACCAAGATGCCA 480
QY 481 TTCTCGTGGAAAGCTGCTGCACTAAATGTTCCGCGGTATTTCTTGATGTCTCTGACAGA 540
DB 481 TTCTCGTGGAAAGCTGCTGCACTAAATGTTCCGCGGTATTTCTTGATGTCTCTGACAGA 540
QY 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGACCTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGACCTGGCGGTGGAGCATC 600
QY 601 TGGTGCATTTGGTTCACGAGCAATCGCGCTGTAGCGGGCCCATTAAGTTCCTGCTCGG 660
DB 601 TGGTGCATTTGGTTCACGAGCAATCGCGCTGTAGCGGGCCCATTAAGTTCCTGCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGAAAGCGCATCGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGAAAGCGCATCGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
DB 781 ATGAGGGCATCGTTCCTCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
QY 841 TGGCGGCCATTACCGAGTCCGCGGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
DB 841 TGGCGGCCATTACCGAGTCCGCGGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
QY 901 AGCATACCGAAGACAGCTCATGTTTATATCCCGCGGTAAACCAATCAAAACAGGATTTTC 960
DB 901 AGCATACCGAAGACAGCTCATGTTTATATCCCGCGGTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACCGAGCTGGAGCGCTTCTGCTCAACTCTCTCAGGCGCCAGGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACCGAGCTGGAGCGCTTCTGCTCAACTCTCTCAGGCGCCAGGGTGA 1020

```


Qy	1021	AGGCAATCAGCTGTTGCCGTCCTCACTGTTGAAAAAGAAAAACAACCTTGGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCGTCCTCACTGTTGAAAAAGAAAAACAACCTTGGCGCCCAATA	1080
Qy	1081	CGAAAAACCGCTCTCCCGCGCGTTGSCCGATTCATTAAATGCAGCTGGCACGACAGGTTT	1140
Db	1081	CGAAAAACCGCTCTCCCGCGCGTTGSCCGATTCATTAAATGCAGCTGGCACGACAGGTTT	1140
Qy	1141	CCGACTGGAAGCGGCGCAGTGACGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG	1200
Db	1141	CCGACTGGAAGCGGCGCAGTGACGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG	1200
Qy	1201	GCACAAATCTCATGTTTGCACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGCACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260
Qy	1261	TCAGGCAACCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTGTAATACACTGCAATAATTCG	1320
Db	1261	TCAGGCAACCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTGTAATACACTGCAATAATTCG	1320
Qy	1321	TGTCGCTCAAGCGCGCACTCCCGTCTCGATAATGTTTTTTCGCCCGACATCATPAACGGTT	1380
Db	1321	TGTCGCTCAAGCGCGCACTCCCGTCTCGATAATGTTTTTTCGCCCGACATCATPAACGGTT	1380
Qy	1381	CTGGCAAAATTTCTGAATCAGCTGTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Db	1381	CTGGCAAAATTTCTGAATCAGCTGTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Qy	1441	ATTGTGAGCGGATAACAATTTTACACAGGAAAAACGCGAGTCGGTTTAGTGTGTTTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAATTTTACACAGGAAAAACGCGAGTCGGTTTAGTGTGTTTTTTCACGA	1500
Qy	1501	GCACCTTCAACAAAGGACCATAGCATATGAAAAAATTAATTCGCAATTCCTTTAGTGTG	1560
Db	1501	GCACCTTCAACAAAGGACCATAGCATATGAAAAAATTAATTCGCAATTCCTTTAGTGTG	1560
Qy	1525	-----AATTAGAAACTGMAAGAGTAAACTGTTAACTCTCG	1560
Db	1561	GTACTTTCTACTCGCGCGATATGAAAACTGGAAGAGTAACTGTTAACTCTCG	1620
Qy	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620
Db	1621	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1680
Qy	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGAAGAGAAATTCACACAGGTT	1680
Db	1681	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGAAGAGAAATTCACACAGGTT	1740
Qy	1681	CGCGCAACTGGGATGCGCTGACATTTCTTGGGCACACGACCGCTTTGTCGCTAC	1740
Db	1741	CGCGCAACTGGGATGCGCTGACATTTCTTGGGCACACGACCGCTTTGTCGCTAC	1800
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800
Db	1801	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1860
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTACAAAGGCAAGCTGATGTTGTTACCCGATCGCTGTT	1860
Db	1861	CCGTTTACCTGGGATGCGGTACGTTACAAAGGCAAGCTGATGTTGTTACCCGATCGCTGTT	1920
Qy	1861	GAACGGTTATCGCTGATTTATAACAAAGATCTGTCGCGAACCCGCAAAAACCTCGGAA	1920
Db	1921	GAACGGTTATCGCTGATTTATAACAAAGATCTGTCGCGAACCCGCAAAAACCTCGGAA	1980
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC	1980
Db	1981	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC	2040
Qy	1981	CTGCAAGAAACGTAATTCACTGCGCCGCTGATGTCGCTGACGGGGTTATGCGTTCAAG	2040
Db	2041	CTGCAAGAAACGTAATTCACTGCGCCGCTGATGTCGCTGACGGGGTTATGCGTTCAAG	2100

Search completed: August 1, 2005, 17:22:05
Job time : 9284 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgacacatcgatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3300	100.0	3300	6	AX202426 Sequence
2	3300	100.0	3300	6	AX202427 Sequence
3	3226.4	97.8	3300	6	AX202430 Sequence
4	3215.2	97.4	3300	6	AX202429 Sequence
5	3207.2	97.2	3300	6	AX202428 Sequence
6	2705.8	82.0	7373	6	AX284144 Sequence
7	2705.8	82.0	8101	6	AX172306 Sequence
8	2701	81.8	6648	6	AX377531 Sequence
9	2700	81.8	7475	6	AX1965 Sequence 1
10	2700	81.8	7475	6	AR031992 Sequence
11	2700	81.8	7475	6	AR207294 Sequence
12	2698	81.8	9191	6	AX377532 Sequence
13	2676.2	81.1	6724	12	AF097412 Sequence
14	2636.8	79.9	6806	6	AX378208 Sequence
15	2631	79.7	6706	12	AF031088 Sequence
16	2598.2	78.7	6748	12	AF031813 Sequence
17	1609.4	48.8	5558	6	CQ794769 Sequence
18	1474	47.0	4700	6	CQ846980 Sequence
19	1462.4	44.3	4921	6	AX768201 Sequence

20	1462.4	44.3	4935	6	AX768200	AX768200 Sequence
21	1462.4	44.3	4945	6	AX768202	AX768202 Sequence
22	1462.4	44.3	4951	6	AX768203	AX768203 Sequence
c	23	1432.4	43.4	6755	12	AF140577 Integrati
	24	1417.6	43.0	1511	6	AX431576 Sequence
	25	1417.6	43.0	4297	12	AX219685 Expressio
	26	1417.6	43.0	4359	12	AX219686 Expressio
	27	1417.6	43.0	4419	12	AX219688 Expressio
	28	1417.6	43.0	4855	12	AX219689 Expressio
	29	1417.6	43.0	4968	12	AX219687 Expressio
	30	1417.6	43.0	6956	12	AX219682 Shuttle e
	31	1417.6	43.0	7018	12	AX219683 Shuttle e
	32	1417.6	43.0	7509	12	AX219684 Shuttle e
	33	1400.8	42.4	4566	12	AF177933 Cloning v
	34	1400.8	42.4	4635	12	AF177932 Cloning v
	35	1400.8	42.4	6399	12	AX302759 Expressio
	36	1400.8	42.4	6587	12	AX289542 Expressio
	37	1362	41.3	5201	6	AR493834 Sequence
	38	1362	41.3	5201	6	AR493835 Sequence
	39	1329.2	40.3	4806	12	AF050464 Expressio
c	40	1238.4	37.5	4803	12	AY243506 Cloning v
c	41	1238.4	37.5	5024	6	CQ846977 Sequence
	42	1237.2	37.5	1922	6	AX478528 Sequence
	43	1204	36.5	4506	12	AB116367 Expressio
	44	1204	36.5	4557	6	E13740 DNA sequenc
	45	1204	36.5	4557	6	AR373251 Sequence

ALIGNMENTS

RESULT 1
AX202426
LOCUS AX202426 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 14 from Patent WO0152620.
ACCESSION AX202426
VERSION AX202426.1 GI:15392173
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLES Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 14 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-ml and zinc finger protein 2FPml"

ORIGIN	Query Match	100.0%;	Score 3300;	DB 6;	Length 3300;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3300;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	CCGACACCATCGAATGGTGC	AAACCTTTCCGCGGTATGCGCATGATGAGCCCGGAAGAGA	60	
Db	1	CCGACACCATCGAATGGTGC	AAACCTTTCCGCGGTATGCGCATGATGAGCCCGGAAGAGA	60	
QY	61	GTCAATTCAGGTGGTGAATG	TGAACCACTTATACGATGTCGAGATGCGG	120	
Db	61	GTCAATTCAGGTGGTGAATG	TGAACCACTTATACGATGTCGAGATGCGG	120	
QY	121	GTGTCCTTATCAGACCGG	TTTCCCGCGGTGGTGAACCACTTCTGCGGAAAA	180	
Db	121	GTGTCCTTATCAGACCGG	TTTCCCGCGGTGGTGAACCACTTCTGCGGAAAA	180	
QY	181	CGCGGAAAAAGTGGAAGCG	CGCATGGCGAGCTGAATTACATCCCAACCGCGTGGCAC	240	

Db	181		CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATACATTCCTCAACCGCGTGGCAC	240
QY	241		AACAACTGGGGGCAAAAGTCTGCTGATTTGGCGTTCGCCACCTCCAGTCTGGCCCTGC	300
Db	241		AACAACTGGCGGGCAAAAGTCTGCTGATTTGGCGTTCGCCACCTCCAGTCTGGCCCTGC	300
QY	301		ACGCGCGCTCGCAAAATTTGTCGGGCGATTAATCTCGCGCGCATCAACTGGGTGCCAGG	360
Db	301		ACGCGCGCTCGCAAAATTTGTCGGGCGATTAATCTCGCGCGCATCAACTGGGTGCCAGG	360
QY	361		TGGTGGTTCGATGGTAGAACGAAGCGCGTTCGAAGCCCTGTAAAGCGCGGTGCACAAATC	420
Db	361		TGGTGGTTCGATGGTAGAACGAAGCGCGTTCGAAGCCCTGTAAAGCGCGGTGCACAAATC	420
QY	421		TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA	480
Db	421		TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
QY	481		TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA	540
Db	481		TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA	540
QY	541		CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC	600
Db	541		CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC	600
QY	601		TGGTGCATTTGGTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAATGTTCTGTCGG	660
Db	601		TGGTGCATTTGGTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAATGTTCTGTCGG	660
QY	661		CGCGTCTGGCTGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661		CGCGTCTGGCTGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721		CGGAAACGGGAACGGCACTGGAGTGCCATGTCGGTTTTCAACAAACCAATGCAATGCTGA	780
Db	721		CGGAAACGGGAACGGCACTGGAGTGCCATGTCGGTTTTCAACAAACCAATGCAATGCTGA	780
QY	781		ATGAGGGATCGTCCCACTGGATGCTGTTGCCAAGCATCAGATGCGCTGGCGGCA	840
Db	781		ATGAGGGATCGTCCCACTGGATGCTGTTGCCAAGCATCAGATGCGCTGGCGGCA	840
QY	841		TGCGGCCATTAACGAGTCCGGCTGCGGCTGCGGATATCTCGGTAGTGGGATACG	900
Db	841		TGCGGCCATTAACGAGTCCGGCTGCGGCTGCGGATATCTCGGTAGTGGGATACG	900
QY	901		ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCAACATCAAAACAGGATTTTC	960
Db	901		ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCAACATCAAAACAGGATTTTC	960
QY	961		GCCTGCTGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA	1020
Db	961		GCCTGCTGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA	1020
QY	1021		AGGCAATCAGCTGTTGCCGCTCACTGGTGAAGAAAAACCAACCTGGCGGCCAATA	1080
Db	1021		AGGCAATCAGCTGTTGCCGCTCACTGGTGAAGAAAAACCAACCTGGCGGCCAATA	1080
QY	1081		CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGCGACAGAGTTTTC	1140
Db	1081		CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGCGACAGAGTTTTC	1140
QY	1141		CCCGACTGGAAGCGGGCAGTGAGCGCAACCGCAATTAATGTAGTTAGCTCACCTATTAG	1200
Db	1141		CCCGACTGGAAGCGGGCAGTGAGCGCAACCGCAATTAATGTAGTTAGCTCACCTATTAG	1200
QY	1201		GCACAAATCTCATGTTTGACAGCTTATCATCGATGCAAGTGCACCAATCTTCTGCGC	1260
Db	1201		GCACAAATCTCATGTTTGACAGCTTATCATCGATGCAAGTGCACCAATCTTCTGCGC	1260
QY	1261		TCAGGACCCATCGGAAGCTGTGGTATGCGTGCAGGTGCTGAATCACTGCATTAATTCG	1320
Db	1261		TCAGGACCCATCGGAAGCTGTGGTATGCGTGCAGGTGCTGAATCACTGCATTAATTCG	1320

Db	1261		TCAGGACCCATCGGAAGCTGTGGTATGCGTGTGCGAGGTCGTAAATCACTGCATTAATTCG	1320
QY	1321		TGTCGCTCAAGCGCGCACTCCCGTTCTGGATTAATGTTTTTTCGGCGGACATCATAAACGGTT	1380
Db	1321		TGTCGCTCAAGCGCGCACTCCCGTTCTGGATTAATGTTTTTTCGGCGGACATCATAAACGGTT	1380
QY	1381		CTGGCAAAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTTGGA	1440
Db	1381		CTGGCAAAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTTGGA	1440
QY	1441		ATTGTGACGCGATTAACAAATTTTCACAGGAAACAGCCAGTCCGTTTATAGGTTTTCACGA	1500
Db	1441		ATTGTGACGCGATTAACAAATTTTCACAGGAAACAGCCAGTCCGTTTATAGGTTTTCACGA	1500
QY	1501		GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTCGTAAATCTGG	1560
Db	1501		GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTCGTAAATCTGG	1560
QY	1561		ATTAAACGCGATTAAGGCTATAACGGTCTCGCTGAAGTCGCTGAAGAAATTCGAGAAAGAT	1620
Db	1561		ATTAAACGCGATTAAGGCTATAACGGTCTCGCTGAAGTCGCTGAAGAAATTCGAGAAAGAT	1620
QY	1621		ACCGAAATTAAGTCAACCGTTGAGCATCCGATATACTGGAAGAGAAATTCGCCACAGTT	1680
Db	1621		ACCGAAATTAAGTCAACCGTTGAGCATCCGATATACTGGAAGAGAAATTCGCCACAGTT	1680
QY	1681		GGCGAACTGGCGATGCGCTGACATTTCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681		GGCGAACTGGCGATGCGCTGACATTTCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741		GCTCAATCTGCGCTGTTGGCTGAAATCACCCCGACAAAGGTTTCCAGGACAAGCTGTAT	1800
Db	1741		GCTCAATCTGCGCTGTTGGCTGAAATCACCCCGACAAAGGTTTCCAGGACAAGCTGTAT	1800
QY	1801		CCGTTTACCTGGGATGCGCTACGTTACAAACGGCAAGCTGATTTGCTTACC CGATCGCTGTT	1860
Db	1801		CCGTTTACCTGGGATGCGCTACGTTACAAACGGCAAGCTGATTTGCTTACC CGATCGCTGTT	1860
QY	1861		GAACGTTATGCTGATTTATAACAAAGATCTGCTGCGCAACCGCCGCAAAACCTGGGAA	1920
Db	1861		GAACGTTATGCTGATTTATAACAAAGATCTGCTGCGCAACCGCCGCAAAACCTGGGAA	1920
QY	1921		GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCTGATGTTTCAAC	1980
Db	1921		GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCTGATGTTTCAAC	1980
QY	1981		CTGCAAGAACCGTATTTCACTGCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG	2040
Db	1981		CTGCAAGAACCGTATTTCACTGCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG	2040
QY	2041		TATGAAACCGCAAGTACGACATTAAGAGCTGGGCTGGATTAACGCTGGCGGGAAGCG	2100
Db	2041		TATGAAACCGCAAGTACGACATTAAGAGCTGGGCTGGATTAACGCTGGCGGGAAGCG	2100
QY	2101		GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACACATGATGACAGACACCGATTAC	2160
Db	2101		GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACACATGATGACAGACACCGATTAC	2160
QY	2161		TCCATCGAAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACCGGCCGCTGG	2220
Db	2161		TCCATCGAAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACCGGCCGCTGG	2220
QY	2221		GCATGTCACCAATCGACACCGGCTTCTGTTGGCGTCTGAGCGCAGGTATTAACCGCCAGT	2280
Db	2221		GCATGTCACCAATCGACACCGGCTTCTGTTGGCGTCTGAGCGCAGGTATTAACCGCCAGT	2280
QY	2281		AAGGTCACCAATCGACACCGGCTTCTGTTGGCGTCTGAGCGCAGGTATTAACCGCCAGT	2340
Db	2281		AAGGTCACCAATCGACACCGGCTTCTGTTGGCGTCTGAGCGCAGGTATTAACCGCCAGT	2340
QY	2341		CCGAAACAAAGAGCTGGCAAAAGGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTTCTG	2400
Db	2341		CCGAAACAAAGAGCTGGCAAAAGGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTTCTG	2400

Qy	2401	GAAGCGGTTTAATAAGACAAACCCTCGGTGGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTTAATAAGACAAACCCTCGGTGGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAGAGTCCACGTAATGGCGCCACCATGGMAAACCGCCAGAAAGGTGAATCATTG	2520
Db	2461	TTGGCGAAGAGTCCACGTAATGGCGCCACCATGGMAAACCGCCAGAAAGGTGAATCATTG	2520
Qy	2521	CCGAAACATCCGCGAGATGTCCGCTTTCTGGTATGCCGTGGCTACTCGCGGTGATCAACGCC	2580
Db	2521	CCGAAACATCCGCGAGATGTCCGCTTTCTGGTATGCCGTGGCTACTCGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGGTCAGACTGTGTAAGACCTGAAAGACCGCGCAGACTAAATCGAGCTCG	2640
Db	2581	GCCAGCGGTGGTCAGACTGTGTAAGACCTGAAAGACCGCGCAGACTAAATCGAGCTCG	2640
Qy	2641	AACAACACCAACATAACATAACAACTCCGGATCGAGGAAGGATTTCAGAAATTC	2700
Db	2641	AACAACACCAACATAACATAACAACTCCGGATCGAGGAAGGATTTCAGAAATTC	2700
Qy	2701	GGATCTCTTCTCTGTGGCCAGCGCGCCCTCGAGCCCGGGAGAAAGCCCTATGCTTGT	2760
Db	2701	GGATCTCTTCTCTGTGGCCAGCGCGCCCTCGAGCCCGGGAGAAAGCCCTATGCTTGT	2760
Qy	2761	CCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGGCCACACGCGTACCCAC	2820
Db	2761	CCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGGCCACACGCGTACCCAC	2820
Qy	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTATGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTATGCCAGTCCAGCAAC	2880
Qy	2881	CTGGTGGCCCATCAACGACCTCATCTACTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC	2940
Db	2881	CTGGTGGCCCATCAACGACCTCATCTACTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC	2940
Qy	2941	AAGTCTTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAAGCTACTCACCGGGGAGAAG	3000
Db	2941	AAGTCTTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAAGCTACTCACCGGGGAGAAG	3000
Qy	3001	CCCTATGCTTGTCCGGAAATGTGTAAGTCTTTCAGCCGACGAGTAACTCGTGGCGCCAC	3060
Db	3001	CCCTATGCTTGTCCGGAAATGTGTAAGTCTTTCAGCCGACGAGTAACTCGTGGCGCCAC	3060
Qy	3061	CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTCCGCGCAATCTTTTAGC	3120
Db	3061	CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTCCGCGCAATCTTTTAGC	3120
Qy	3121	CAGCGCGCCACCTGGCCAGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGT	3180
Db	3121	CAGCGCGCCACCTGGCCAGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGT	3180
Qy	3181	CCAGAAATGTGGCAAGTCTTTCTCTCGGTGTGACAAATCTCGTCCGGCACCAACGTAATC	3240
Db	3181	CCAGAAATGTGGCAAGTCTTTCTCTCGGTGTGACAAATCTCGTCCGGCACCAACGTAATC	3240
Qy	3241	ACCGGTAAAAAATAGTGGCCAGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT	3300
Db	3241	ACCGGTAAAAAATAGTGGCCAGCGCGCCAGTACCGGTACGACGTTCCGGACTACGCT	3300
RESULT 2			
AX202427			
LOCUS	AX202427	3300 bp	DNA
DEFINITION	Sequence 15 from Patent WO0152620.		linear
ACCESSION	AX202427		
VERSION	AX202427.1	GI:15392175	
KEYWORDS			
SOURCE			synthetic construct
ORGANISM			synthetic construct
			other sequences; artificial sequences.
REFERENCE 1			

841 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
Db
841 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAAGGATTTTC 960
Qy 961 GCGTGTGGGGCAAAACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCGTGTGGGGCAAAACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCCGTCTCACCTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCACCTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAAGAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAAGAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 1200
Qy 1201 GCACAATTCATGTTGACGCTTATCATGACTGCACGCTGCAACCAATGTTCTGGCG 1260
Db 1201 GCACAATTCATGTTGACGCTTATCATGACTGCACGCTGCAACCAATGTTCTGGCG 1260
Qy 1261 TCAGCGCCATCGAAGCTGCGTATGCGTGTGAGCTGTAATCATGCTGATTAATTCG 1320
Db 1261 TCAGCGCCATCGAAGCTGCGTATGCGTGTGAGCTGTAATCATGCTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGCTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGCTT 1380
Qy 1381 CTGGCAATTTCTGAATGAGCTGTTGACATTAATCATCGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAATTTCTGAATGAGCTGTTGACATTAATCATCGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTGCGGATACAAATTTACACAGGAACAGCCAGTCCGTTAGTGTTCACGA 1500
Db 1441 ATTGTGCGGATACAAATTTACACAGGAACAGCCAGTCCGTTAGTGTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGTAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGTAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGGAAGAAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGGAAGAAGCTGAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTACGTTTACACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTACACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTAACAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTAACAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920

1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAAACCGTATCTTCACTGCGCGCTGATGCTGCTGACGGGGTATTCGCTTCAAG 2040
Db 1981 CTGCAAGAAACCGTATCTTCACTGCGCGCTGATGCTGCTGACGGGGTATTCGCTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAAACAAACACATGAATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAAACAAACACATGAATGACAGACCGGATTAC 2160
Qy 2161 TCCATCGCAGAGTGGCTTTAATAAAGCGAAGACAGCGATGACATCAACGGCGCGTGG 2220
Db 2161 TCCATCGCAGAGTGGCTTTAATAAAGCGAAGACAGCGATGACATCAACGGCGCGTGG 2220
Qy 2221 GCATGGTCCACATCGACACGACCAAGTGAATTTATGGTGTAAACGGTACTGCGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACGACCAAGTGAATTTATGGTGTAAACGGTACTGCGACCTTC 2280
Qy 2281 AAGGCTCAACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGCTCAACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
Qy 2341 CCGAAACAAAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CCGAAACAAAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
Qy 2401 GAACGGTTTAATAAAGACAAACCGCTGGTGGCTGGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAACGGTTTAATAAAGACAAACCGCTGGTGGCTGGCTGAAAGTCTTACGAGGAAG 2460
Qy 2461 TTGCGAAAGATCCACGCTATTGCGCCACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGAAAGATCCACGCTATTGCGCCACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGGCTTTCTGATGTCGCTGGCTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGGCTTTCTGATGTCGCTGGCTACTGCGGTGATCAAGCC 2580
Qy 2581 GCCAGCGTCTGACAGTGTGATGAAGCCTGAAAGACGCGAGACTTAATTCAGGCTG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCTGAAAGACGCGAGACTTAATTCAGGCTG 2640
Qy 2641 AACCAACAAATAAACAACAACTCGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACCAACAAATAAACAACAACTCGGATCGAGGGAAGGATTTCAAGATTC 2700
Qy 2701 GGATCTCTTCTGTCGGCGGCGCTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCTCTTCTGTCGGCGGCGCTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTTGTAAGTCTTCTCTCAGAGTCTCACTGGTGGCCACCGAGCTACCCAC 2820
Db 2761 CCGGAATGTTGTAAGTCTTCTCTCAGAGTCTCACTGGTGGCCACCGAGCTACCCAC 2820
Qy 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTTAGCAGTCCAGCAAC 2880
Db 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTTAGCAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGCTCATCTGCGGAGAGCATACATCTCCAGATGTTGGC 2940
Db 2881 CTGGTGGCCATCAACGCTCATCTGCGGAGAGCATACATCTCCAGATGTTGGC 2940
Qy 2941 AAGTCTTTCTCTCGGTGACAAATCTCGTCCGGACCAAGTACTCACACCGGGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTGACAAATCTCGTCCGGACCAAGTACTCACACCGGGAAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTACGCGCAGGATAACTGTTGGTGGCCAC 3060

```

Db      3001  |||||CCCTATGCTGTGTCGGAATAGTGGTAAGTCTTTCAGCGCAGCGGATAAACCCTGCTGGCCAC 3060
Qy      3061  CAGCGGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db      3061  CAGCGGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Qy      3121  CAGCGCGGCCACTCGGCCAGCCATCAAGCAGCTCATACCTGAGAGAGCCATACAAATGT 3180
Db      3121  CAGCGCGGCCACTCGGCCAGCCATCAAGCAGCTCATACCTGAGAGAGCCATACAAATGT 3180
Qy      3181  CCAGAAATGGCAAGCTCTTCTCGGTCTGACAACTCGTCCGCGACCAAGTACTCAC 3240
Db      3181  CCAGAAATGGCAAGCTCTTCTCGGTCTGACAACTCGTCCGCGACCAAGTACTCAC 3240
Qy      3241  ACCCGTAAAAAACTAGTGGCCAGGCCCGCCAGTACCCCGTACGAGCTTCGCGACTACGCT 3300
Db      3241  ACCCGTAAAAAACTAGTGGCCAGGCCCGCCAGTACCCCGTACGAGCTTCGCGACTACGCT 3300

RESULT 3
LOCUS   AX202430 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 18 from Patent WO0152620.
ACCESSION AX202430
VERSION AX202430.1 GI:15392178
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 18 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)

FEATURES
source
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-Ap3 and zinc finger protein
ZFPAP3"

ORIGIN
Query Match 97.8%; Score 3226.4; DB 6; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0; Mismatches 46; Indels 0; Gaps 0;
Matches 3254; Conservative 0;

Qy      1 CGACACCATCGAATGGTGCAAAACCTTTGCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db      1 CGACACCATCGAATGGTGCAAAACCTTTGCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy      61 GTCAATTACGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCG 120
Db      61 GTCAATTACGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCG 120
Qy      121 GGTCTCTTATCAGACCGTTTCCCGGTGTGAACAGCGCCAGCCAGCTTTCTGCGAATA 180
Db      121 GGTCTCTTATCAGACCGTTTCCCGGTGTGAACAGCGCCAGCCAGCTTTCTGCGAATA 180
Qy      181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTAACATTCCTCAACCGCGTGGCAC 240
Db      181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTAACATTCCTCAACCGCGTGGCAC 240
Qy      241 AACAACTGCGGGCAACAGTCTGTTGCTGATTGGGTTGCCACCTCAGTCTGGCCCTGC 300
Db      241 AACAACTGCGGGCAACAGTCTGTTGCTGATTGGGTTGCCACCTCAGTCTGGCCCTGC 300
Qy      301 ACGCCCGTTCGCAAAATGTTCGGCGGATTAATCTCGGCCGATCAACTGGGTGCCAGCG 360
Db      301 ACGCCCGTTCGCAAAATGTTCGGCGGATTAATCTCGGCCGATCAACTGGGTGCCAGCG 360

361 TGTGTGTCGATCGATGGTAGAACGAAAGCGCGCTGCGAAGCCTGTAAAGCGCGGTGCAATC 420
361 TGTGTGTCGATCGATGGTAGAACGAAAGCGCGCTGCGAAGCCTGTAAAGCGCGGTGCAATC 420
421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATATCCGCTGGATCAGCAGGATGCCA 480
421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATATCCGCTGGATCAGCAGGATGCCA 480
481 TTGTCTGTGGAAGCTCGCTGCACTAATGTTCCGGCGTATTTCTTGTGATGCTCTGACCA 540
481 TTGTCTGTGGAAGCTCGCTGCACTAATGTTCCGGCGTATTTCTTGTGATGCTCTGACCA 540
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGTACGCGATCGGCGTGGAGCATC 600
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGTACGCGATCGGCGTGGAGCATC 600
601 TGTGTGCAATTCGGTCCAGCAAAATCGCGCTGTTAGCGGCGCCATTAAGTTCTGTCTCG 660
601 TGTGTGCAATTCGGTCCAGCAAAATCGCGCTGTTAGCGGCGCCATTAAGTTCTGTCTCG 660
661 CGCGTCTCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
661 CGCGTCTCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
721 CGGAACGGGAAGCGGACATGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
721 CGGAACGGGAAGCGGACATGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
781 ATGAGGCGCATGTTCCCACTCGGATGCTGTTGCAACAGATCAGATGCGCTGGCGGCA 840
781 ATGAGGCGCATGTTCCCACTCGGATGCTGTTGCAACAGATCAGATGCGCTGGCGGCA 840
841 TGC CGCGCCATTAACGAGTCCCGGCTGCGGTGTTGTCGGATATCTCGGTAGTGGGATCG 900
841 TGC CGCGCCATTAACGAGTCCCGGCTGCGGTGTTGTCGGATATCTCGGTAGTGGGATCG 900
901 ACGATACGGAAGACAGTCTATGTTATATCCCGCTTAAACCAATCAAAACAGGATTTTC 960
901 ACGATACGGAAGACAGTCTATGTTATATCCCGCTTAAACCAATCAAAACAGGATTTTC 960
961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
1021 AGGCAATCAGCTGTGTGCCCGTCTCACTGTGTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
1021 AGGCAATCAGCTGTGTGCCCGTCTCACTGTGTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCGCTGGCAGCAGAGTTT 1140
1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCGCTGGCAGCAGAGTTT 1140
1141 CCCGACTGGAAGCGGCGGATGAGCGCAACCGCAATTAATGTAGTTAGTCTCACTATTAG 1200
1141 CCCGACTGGAAGCGGCGGATGAGCGCAACCGCAATTAATGTAGTTAGTCTCACTATTAG 1200
1201 GCACAATTTCTCATGTTTGACAGCTTATCATGCACTGCGGTGACCAATGCTTCTGGCG 1260
1201 GCACAATTTCTCATGTTTGACAGCTTATCATGCACTGCGGTGACCAATGCTTCTGGCG 1260
1261 TCAGGACGCCATTCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATAATTCG 1320
1261 TCAGGACGCCATTCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATAATTCG 1320
1321 TGTCTCTCAAGCGGCACTCCCGTTCGGATTAATGTTTTTGGCGCGACATCATACGGTT 1380
1321 TGTCTCTCAAGCGGCACTCCCGTTCGGATTAATGTTTTTGGCGCGACATCATACGGTT 1380
1381 CTGGCAAAATTTCTCAAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTATATGTGTGA 1440
1381 CTGGCAAAATTTCTCAAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTATATGTGTGA 1440
1441 ATTGTGAGCGGATAACAAATTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
```


Db 1441 ATTGTGACGGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
QY GCACCTTACCAACAAAGGACCATAGATTATGAAACTGGAAGAGGTAAACTGTTATCTGG 1560
Db GCACCTTACCAACAAAGGACCATAGATTATGAAACTGGAAGAGGTAAACTGTTATCTGG 1560
QY ATTAACGGCGATAAAGGCTATTAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db ATTAACGGCGATAAAGGCTATTAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY ACCGGAATTAAGTACACCGTTGAGCATCCGATGAACTGGAAGAGAAATTCAGAGGTT 1680
Db ACCGGAATTAAGTACACCGTTGAGCATCCGATGAACTGGAAGAGAAATTCAGAGGTT 1680
QY GCGGCAATCGGCGATGGCCCTGACATATCTTCTGGCACACGACCGCTTTTGGTGCTAC 1740
Db GCGGCAATCGGCGATGGCCCTGACATATCTTCTGGCACACGACCGCTTTTGGTGCTAC 1740
QY GCTCAATCTGGCCTGTTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db GCTCAATCTGGCCTGTTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY CCGTTTACCTGGGATGCGGTAGCTTTACACCGCAAGCTGATGCTTACCGCATGCTGTT 1860
Db CCGTTTACCTGGGATGCGGTAGCTTTACACCGCAAGCTGATGCTTACCGCATGCTGTT 1860
QY GAAGCGTTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
Db GAAGCGTTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
QY GAGATCCCGGCGCTGGATAAAGAACTGAAAGCAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db GAGATCCCGGCGCTGGATAAAGAACTGAAAGCAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY CTGCAAGAACGCTACTTACCTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db CTGCAAGAACGCTACTTACCTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
Db TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
QY GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACCAACATGAATGACAGACACCGATTAC 2160
Db GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACCAACATGAATGACAGACACCGATTAC 2160
QY TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY GCATGGTCCAACTCGACACAGCAAAAGTGAATTAATGGTAAACGGTACTCGCGACCTTC 2280
Db GCATGGTCCAACTCGACACAGCAAAAGTGAATTAATGGTAAACGGTACTCGCGACCTTC 2280
QY AAGGTCAACCATCAAAACCGTTCGTTGGCTGCTGAGCGAGGTATTAAACCGCGCCAGT 2340
Db AAGGTCAACCATCAAAACCGTTCGTTGGCTGCTGAGCGAGGTATTAAACCGCGCCAGT 2340
QY CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400
Db CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400
QY GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGAGCGCTGAAGTCTTACGAGGAAG 2460
Db GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGAGCGCTGAAGTCTTACGAGGAAG 2460
QY TTGGCGAAAGATCCACGATTATGCGGCCACCAATGGAAGAAACGCCCAAGAAAGGTGAATCATG 2520
Db TTGGCGAAAGATCCACGATTATGCGGCCACCAATGGAAGAAACGCCCAAGAAAGGTGAATCATG 2520
QY CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGCGGTGCGTACTGCGGTGATCAACGCC 2580

Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCCGCTGGGTACTGCGGTGATCAACGCC 2580
QY GCCAGCGTCTGACAGTCTGATGATGAAGCCTGAAAGACGCGAGACTTAATTCGAGCTCG 2640
Db GCCAGCGTCTGACAGTCTGATGATGAAGCCTGAAAGACGCGAGACTTAATTCGAGCTCG 2640
QY AACCAACCAACAATAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAATTC 2700
Db AACCAACCAACAATAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAATTC 2700
QY GGATCCTCTCTCTCTGTCGCGCCAGCGCGCTTCAGCGCGGGGAGAGCCCTATGCTGT 2760
Db GGATCCTCTCTCTCTGTCGCGCCAGCGCGCTTCAGCGCGGGGAGAGCCCTATGCTGT 2760
QY CCGGAATGTGTAGTCTCTCTCAGAGCTCTCAGCTGGTGGCCACCGAGCGTACCCAC 2820
Db CCGGAATGTGTAGTCTCTCTCAGAGCTCTCAGCTGGTGGCCACCGAGCGTACCCAC 2820
QY ACCGGTGAACCAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCCAGTCCAGCAAC 2880
Db ACCGGTGAACCAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCCAGTCCAGCAAC 2880
QY CTGCTGGCCATCAACCGCTATCTGCGGAGAGGATCAACAAATGTCAGAAATGTGGC 2940
Db CTGCTGGCCATCAACCGCTATCTGCGGAGAGGATCAACAAATGTCAGAAATGTGGC 2940
QY AAGTCTTTCTCTCGGTCTGCAATCTCGTCCGCAACCAAGTACTCACACGGGAGAG 3000
Db AAGTCTTTCTCGGTCTGCAATCTCGTCCGCAACCAAGTACTCACACGGGAGAG 3000
QY CCTATGCTTTCTCGGAAATGTGTAAGTCTTCTAGCAGCAGTGGCTCTTTAGTAC 3060
Db CCTATGCTTTCTCGGAAATGTGTAAGTCTTCTAGCAGCAGTGGCTCTTTAGTAC 3060
QY CAGGTACCCACACGGGTGAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGC 3120
Db CAGGTACCCACACGGGTGAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGC 3120
QY CAGCGCGCCACCTTGGCCAGCATCAACGCTCATCTGCGGAGAGGCAATACAAATGT 3180
Db CAGCGCGCCACCTTGGCCAGCATCAACGCTCATCTGCGGAGAGGCAATACAAATGT 3180
QY CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGCAACCAAGTACTAC 3240
Db CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGCAACCAAGTACTAC 3240
QY ACCGTAATAAANAACCTAGTGGCGCGCGCGTACCGGTACGAGTTCGGGACTACGT 3300
Db ACCGTAATAAANAACCTAGTGGCGCGCGCGTACCGGTACGAGTTCGGGACTACGT 3300

RESULT 4

AX202429 LOCUS AX202429 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 17 from Patent WO0152620.
ACCESSION AX202429
VERSION AX202429.1 GI:15392177

KEYWORDS synthetic construct

SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE

1 Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.

AUTHORS Methods and compositions to modulate expression in plants

TITLE Patent: WO 0152620-A 17 26-JUL-2001;

JOURNAL The Scripps Research Institute (US); SYNGENTA AGRICULTURAL

DISCOVERY, INC. (CA)

FEATURES Location/Qualifiers

source

1. .3300

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Partial sequence of pMal-m4 and zinc finger protein"

ORIGIN		2FPm4"		Db			
Query Match		97.4%; Score 3215.2; DB 6; Length 3300;		QY			
Best Local Similarity		98.4%; Pred. No. 0;		QY			
Matches 3247; Conservative		0; Mismatches 53; Indels 0; Gaps 0;		Db			
QY	1	CGGACACCATCGAATGGTGC	60	QY	961	GCCTGCTGGGGCAAAAC	1020
Db	1	CGGACACCATCGAATGGTGC	60	Db	1021	AGGGCAATCAGCTGTTG	1080
QY	61	GTCAATTCAGGGTGGTGA	120	QY	1021	AGGGCAATCAGCTGTTG	1080
Db	61	GTCAATTCAGGGTGGTGA	120	Db	1081	CGCAAAACCGCTCTCT	1140
QY	121	GTGTCCTTTATCAGAC	180	QY	1081	CGCAAAACCGCTCTCT	1140
Db	121	GTGTCCTTTATCAGAC	180	Db	1141	CCGACTCGAAAAAGCG	1200
QY	181	CGCGGAAAAAGTGAAG	240	QY	1141	CCGACTCGAAAAAGCG	1200
Db	181	CGCGGAAAAAGTGAAG	240	Db	1201	GCACAATTCATGTTTG	1260
QY	241	AACAACTGCGGGCAAA	300	QY	1201	GCACAATTCATGTTTG	1260
Db	241	AACAACTGCGGGCAAA	300	Db	1261	TCAGGACGCAATCGGA	1320
QY	301	ACGCGCGTTCGCAAA	360	QY	1261	TCAGGACGCAATCGGA	1320
Db	301	ACGCGCGTTCGCAAA	360	Db	1321	TGTGCTCAAGCGCACT	1380
QY	361	TGGTGGTGTGATGTAG	420	QY	1321	TGTGCTCAAGCGCACT	1380
Db	361	TGGTGGTGTGATGTAG	420	Db	1381	CTGGCAAAATTTCTGA	1440
QY	421	TTCTCGCGCAACGCTC	480	QY	1381	CTGGCAAAATTTCTGA	1440
Db	421	TTCTCGCGCAACGCTC	480	Db	1441	ATTGTAGCGGATAACA	1500
QY	481	TTGCTGTGGAAGCTGC	540	QY	1441	ATTGTAGCGGATAACA	1500
Db	481	TTGCTGTGGAAGCTGC	540	Db	1501	GCACCTTACCAACAA	1560
QY	541	CACCCATCAACAGTAT	600	QY	1501	GCACCTTACCAACAA	1560
Db	541	CACCCATCAACAGTAT	600	Db	1561	ATTAACCGCGATAA	1620
QY	601	TGGTGGCATTTGGTCA	660	QY	1561	ATTAACCGCGATAA	1620
Db	601	TGGTGGCATTTGGTCA	660	Db	1621	ACCGGAATTAAGTCA	1680
QY	661	CGCGTCTGGCTGGCTG	720	QY	1621	ACCGGAATTAAGTCA	1680
Db	661	CGCGTCTGGCTGGCTG	720	Db	1681	GGCGCAACTGGCGAT	1740
QY	721	CGGAACGGGAAGCGCA	780	QY	1681	GGCGCAACTGGCGAT	1740
Db	721	CGGAACGGGAAGCGCA	780	Db	1741	GCTCAATCTGGCCTG	1800
QY	781	ATGAGGGCATCGTTCC	840	QY	1741	GCTCAATCTGGCCTG	1800
Db	781	ATGAGGGCATCGTTCC	840	Db	1801	CGGTTTACCTGGGAT	1860
QY	841	TGCGGCCATTAACGAG	900	QY	1801	CGGTTTACCTGGGAT	1860
Db	841	TGCGGCCATTAACGAG	900	Db	1861	GAAGCGTTATCGCTG	1920
QY	901	ACGATACCGAAGACAG	960	QY	1861	GAAGCGTTATCGCTG	1920
Db	901	ACGATACCGAAGACAG	960	Db	1921	GAGATCCCGGCGCTG	1980
QY	961	GCCTGCTGGGGCAAC	1020	QY	1921	GAGATCCCGGCGCTG	1980
Db				Db	1981	CTGCAAGAACCGTACT	2040
				QY	1981	CTGCAAGAACCGTACT	2040
				Db	2041	TATGAAAAACGCAACT	2100
				QY	2041	TATGAAAAACGCAACT	2100

Qy	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATCGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATCGAGACACCGATTAC	2160
Qy	2161	TCCATCGCAGAAGCTGCTTTTAATAAAGGGGAAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAGAAGCTGCTTTTAATAAAGGGGAAAACAGCGATGACCATCAACGGCCCGTGG	2220
Qy	2221	GCATGTGTCCAAACATCGACACAGCAAAAGTGAATTTATCGGTGTAAGGTACTGCCGACCTTC	2280
Db	2221	GCATGTGTCCAACATCGACACAGCAAAAGTGAATTTATCGGTGTAAGGTACTGCCGACCTTC	2280
Qy	2281	AAGGGTCAACCATCAAAACCGTTCTGTTGGGTGCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Db	2281	AAGGGTCAACCATCAAAACCGTTCTGTTGGGTGCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACCTATCTGCTGACTGATGAAGGTCTG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACCTATCTGCTGACTGATGAAGGTCTG	2400
Qy	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAAGATCCACGTATTTGGCGGCACCATGGAAAACGCCAGAAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTATTTGGCGGCACCATGGAAAACGCCAGAAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCGCGAGATGTCGGCTTTCTGCTGATGCGGTGCTACTCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCGCGAGATGTCGGCTTTCTGCTGATGCGGTGCTACTCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAACAAACAAATACAAATACAAACACTCGGATCGAGGAAGATTTTCAGAAATTC	2700
Db	2641	AACAACAAACAAATACAAATACAAACACTCGGATCGAGGAAGATTTTCAGAAATTC	2700
Qy	2701	GGATCTCTTCTCTGTGGCCGAGGGCGCCCTCGAGCCGGGAGAAAGCCCTATGCTGT	2760
Db	2701	GGATCTCTTCTCTGTGGCCGAGGGCGCCCTCGAGCCGGGAGAAAGCCCTATGCTGT	2760
Qy	2761	CCGGATGTGGTAAGTCTCTCTCAGAGCTCTCACTGTGTGGCCACAGCGGTACCCAC	2820
Db	2761	CCGGATGTGGTAAGTCTCTCTCAGCAGAGCAGCTCCCTGTGTGGCCACAGCGGTACCCAC	2820
Qy	2821	ACGGGTGAAAAACCGTATAATCCAGAGTGGCGAAATCTTTTACGACAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAAACCGTATAATCCAGAGTGGCGAAATCTTTTACGACAGAGCAGCAGC	2880
Qy	2881	CTGGTGGCCATCAACGCACTCATATGCGGAGAAAGCCATACAAATGTCCAGAAATGGC	2940
Db	2881	CTGGTGGCCATCAACGCACTCATATGCGGAGAAAGCCATACAAATGTCCAGAAATGGC	2940
Qy	2941	AAGTCTTTCTCTCGGTCTGACATCTCGTCCGACCAACGCTACTCACACCGGGGAGAAG	3000
Db	2941	AAGTCTTTCTGATTTGTCTGTGATCTTTGGAGGCAACCAACGCTACTCACACCGGGGAGAAG	3000
Qy	3001	CCCTATGCTTTGTCGGAAATGGTAAAGTCTCTTCAGCCGCGAGCGATAACTCGTGTGCCAC	3060
Db	3001	CCCTATGCTTTGTCGGAAATGGTAAAGTCTCTTCCTCAGAGCTCTCACCTGTGTGCCAC	3060
Qy	3061	CAGGTAACCCACACAGGGTGAATAACCGTATAATGCCAGAGTGGCGGAAATCTTTTAGC	3120
Db	3061	CAGGTAACCCACACAGGGTGAATAACCGTATAATGCCAGAGTGGCGGAAATCTTTTAGC	3120
Qy	3121	CAGGCCGGCCACTCTGGCCAGCCATCAACGCACTCATACTGGCGAGAAAGCCATACAAATGT	3180
Db	3121	CGCAGCGATAACCTGTGTGGCCATCAACGCACTCATACTGTGGCGAGAAAGCCATACAAATGT	3180

QY	3181	CCGAAATGTGGCAAGTCTTTCTCTCGGTGTGACAATCTGTCGGGCACCAACGTA	3240
Db	3181	CCGAAATGTGGCAAGTCTTTCTCAACTTCAGGGCAATTTGGTCGGTACCAACGTA	3240
QY	3241	ACCGGTAAAAAACTAGTGGCCAGGCGGCCAGTACCCGTACGACGTTCCGGACTACGCT	3300
Db	3241	ACCGGTAAAAAACTAGTGGCCAGGCGGCCAGTACCCGTACGACGTTCCGGACTACGCT	3300
RESULT 5			
LOCUS	AX202428	3300 bp	DNA linear PAT 30-AUG-2001
DEFINITION	Sequence 16 from Patent WO0152620.		
ACCESSION	AX202428		
VERSION	AX202428.1	GI:15392176	
KEYWORDS			
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	1	other sequences; artificial sequences.	
AUTHORS	Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.		
TITLE	Methods and compositions to modulate expression in plants		
JOURNAL	Patent: WO 0152620-A 16 26-JUL-2001; The Scripps Research Institute (US); SYNGENTA AGRICULTURAL DISCOVERY, INC. (CA)		
FEATURES	Location/Qualifiers		
source	1..3300		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Partial sequence of pMal-m3 and zinc finger protein ZFPm3"		

Query Match	97.2%	Score 3207.2	DB 6	Length 3300	
Best Local Similarity	98.2%	Pred. No. 0			
Matches 3242	Conservative	0	Mismatches 58	Indels 0	Gaps 0
Qy	1	CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGCATGATAGCCGCCGGAAGAGA	60		
Db	1	CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGCATGATAGCCGCCGGAAGAGA	60		
Qy	61	GTCAATTCCAGGCTGGTGAAATGTGAAACCAAGTAACTGATACGATGTCGCAGAGTAGTGCCG	120		
Db	61	GTCAATTCCAGGCTGGTGAAATGTGAAACCAAGTAACTGATACGATGTCGCAGAGTAGTGCCG	120		
Qy	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAAACAGGCCAGCCACGTTTCTGCGAAAA	180		
Db	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAAACAGGCCAGCCACGTTTCTGCGAAAA	180		
Qy	181	CGCGGAAAAAAGTGAAGCGGCGATCGCGAGAGCTGAATTACATTCCTCCACCGCGTGGCAC	240		
Db	181	CGCGGAAAAAAGTGAAGCGGCGATCGCGAGAGCTGAATTACATTCCTCCACCGCGTGGCAC	240		
Qy	241	AACAACCTGGCGGCAACACAGTGGTCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC	300		
Db	241	AACAACCTGGCGGCAACACAGTGGTCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC	300		
Qy	301	ACGCGCGTTCGCAATTTGTCGGGCGATTAATCTTCGCGCCCGATCAACTGGGTGGCAGCG	360		
Db	301	ACGCGCGTTCGCAATTTGTCGGGCGATTAATCTTCGCGCCCGATCAACTGGGTGGCAGCG	360		
Qy	361	TGCTGTGTTCGATGGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACATC	420		
Db	361	TGCTGTGTTCGATGGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACATC	420		
Qy	421	TTCTCCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAGGATGCCA	480		
Db	421	TTCTCCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAGGATGCCA	480		
Qy	481	TTGCTGTGAAGCTGCTGCACTAATGTTTCGGGCGTTATTTCTTGATGTCCTGACCCAGA	540		
Db	481	TTGCTGTGAAGCTGCTGCACTAATGTTTCGGGCGTTATTTCTTGATGTCCTGACCCAGA	540		

Qy	541	CACCCATCAA	CAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGAGCATC	600
Db	541	CACCCATCAA	CAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGAGCATC	600
Qy	601	TGCTCGCAT	TGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG	660
Db	601	TGCTCGCAT	TGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG	660
Qy	661	CGCGTCTCG	CTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTCG	CTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAACGG	GAAGGACCTGGAGTGCATGTCGGTTTCAACAACCATGCAATGCTGA	780
Db	721	CGGAACGG	GAAGGACCTGGAGTGCATGTCGGTTTCAACAACCATGCAATGCTGA	780
Qy	781	ATGAGGG	CATCTGCCACTGCGATGCTGTTGCCAACGATCAGATGCGCTGGCGCAA	840
Db	781	ATGAGGG	CATCTGCCACTGCGATGCTGTTGCCAACGATCAGATGCGCTGGCGCAA	840
Qy	841	TGCGGCC	ATTACCGAGTCCGGCTGCGGTGGGTGCGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGGCC	ATTACCGAGTCCGGCTGCGGTGGGTGCGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATAC	CGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAAAACAGGATTTTC	960
Db	901	ACGATAC	CGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAAAACAGGATTTTC	960
Qy	961	GCCTGCT	GGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGGGGTGA	1020
Db	961	GCCTGCT	GGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGGGGTGA	1020
Qy	1021	AGGGCAAT	CAGCTGTGGCCCGTCTCACTGGTGAAAGAAAACCAACCTGGCGCCCAATA	1080
Db	1021	AGGGCAAT	CAGCTGTGGCCCGTCTCACTGGTGAAAGAAAACCAACCTGGCGCCCAATA	1080
Qy	1081	CGCAAA	CCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGCGACGACAGGTTT	1140
Db	1081	CGCAAA	CCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGCGACGACAGGTTT	1140
Qy	1141	CCGACTG	GGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
Db	1141	CCGACTG	GGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
Qy	1201	GCACAAT	TCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
Db	1201	GCACAAT	TCTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
Qy	1261	TCAGCAG	CCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTPAATCACTGCATAAATCG	1320
Db	1261	TCAGCAG	CCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTPAATCACTGCATAAATCG	1320
Qy	1321	TGTCGCT	CANAGCGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGGTT	1380
Db	1321	TGTCGCT	CANAGCGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGGTT	1380
Qy	1381	CTGGCAAA	TATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA	1440
Db	1381	CTGGCAAA	TATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA	1440
Qy	1441	ATTGTAG	CGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCAGGA	1500
Db	1441	ATTGTAG	CGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCAGGA	1500
Qy	1501	GCATTTC	ACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTTAATCTGG	1560
Db	1501	GCATTTC	ACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTTAATCTGG	1560
Qy	1561	ATTAACG	CGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAG	1620
Db	1561	ATTAACG	CGGATAAAGGCTATAAAGCTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAG	1620

Qy	1621	ACCGAATTA	AAGTCAACGGTTGACATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
Db	1621	ACCGAATTA	AAGTCAACGGTTGACATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
Qy	1681	CGGCAACT	GGCGATGGCCCTGCATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGGCAACT	GGCGATGGCGCTGCATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAAT	CTGGCTGTGGCTGGAATCAACCCCGACAAGCGTTCCAGGAACAAGCTGTAT	1800
Db	1741	GCTCAAT	CTGGCTGTGGCTGGAATCAACCCCGACAAGCGTTCCAGGAACAAGCTGTAT	1800
Qy	1801	CGGTTT	ACCTGGGATGCGTACGTTCAACGGCAAGCTGATTGCTTACCCGATCGCTGT	1860
Db	1801	CGGTTT	ACCTGGGATGCGTACGTTCAACGGCAAGCTGATTGCTTACCCGATCGCTGT	1860
Qy	1861	GAACGG	TTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
Db	1861	GAACGG	TTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
Qy	1921	GAGATCC	CGCGCTGGATTAAGAACTGAAACGGAAAGTAAAGAGCGCTGATGTTCAAC	1980
Db	1921	GAGATCC	CGCGCTGGATTAAGAACTGAAACGGAAAGTAAAGAGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAG	AAACCGTATTCACCTGGCGCTGATTCGCTGACGGGGTTATCGCTTCAAG	2040
Db	1981	CTGCAAG	AAACCGTATTCACCTGGCGCTGATTCGCTGACGGGGTTATCGCTTCAAG	2040
Qy	2041	TATGAAA	ACGGCAAGTAAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAA	ACGGCAAGTAAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG	2100
Qy	2101	GGTCTG	ACCTTCTGGTTGACCTGATTAAAAACAACATGAAATGCGACACCGATTAC	2160
Db	2101	GGTCTG	ACCTTCTGGTTGACCTGATTAAAAACAACATGAAATGCGACACCGATTAC	2160
Qy	2161	TCCATCG	CAGAGCTGCTTTAATAAGGCGAAACAGCATGCCATCAACGGCCCGTGG	2220
Db	2161	TCCATCG	CAGAGCTGCTTTAATAAGGCGAAACAGCATGCCATCAACGGCCCGTGG	2220
Qy	2221	GCATGGT	CCAACTCGACCAAGTAAATATGGTGTAAACGGTACTGCGGACCTTC	2280
Db	2221	GCATGGT	CCAACTCGACCAAGTAAATATGGTGTAAACGGTACTGCGGACCTTC	2280
Qy	2281	AAGGGT	CAACCATCAACCGTTCTGGCGTGTGAGCGCAGGTATTAACGCCCCAGT	2340
Db	2281	AAGGGT	CAACCATCAACCGTTCTGGCGTGTGAGCGCAGGTATTAACGCCCCAGT	2340
Qy	2341	CCGAACA	AGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGCTCG	2400
Db	2341	CCGAACA	AGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGCTCG	2400
Qy	2401	GAAGCG	TTAATAAGACAAACCGCTGGGTGCGCTGAACTCTTACGAGGAAGAG	2460
Db	2401	GAAGCG	TTAATAAGACAAACCGCTGGGTGCGCTGAACTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCG	AAAGATTCACGTTATTCGCCCAACCATGGAAGAAACGCCAGAAAGTGAATCATG	2520
Db	2461	TTGGCG	AAAGATTCACGTTATTCGCCCAACCATGGAAGAAACGCCAGAAAGTGAATCATG	2520
Qy	2521	CCGAAC	ATCCCGCAGATGTCGGCTTTCTGGTATGCCGCTGCTACTCGCGTGTATCAACGCC	2580
Db	2521	CCGAAC	ATCCCGCAGATGTCGGCTTTCTGGTATGCCGCTGCTACTCGCGTGTATCAACGCC	2580
Qy	2581	GCCAGCG	TCTGACCTGTAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCG	TCTGACCTGTAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Qy	2641	AACAACA	CAACAATAACAACACCTTCGGGATCGAGGGAAGGATTTTCAAGATTC	2700
Db	2641	AACAACA	CAACAATAACAACACCTTCGGGATCGAGGGAAGGATTTTCAAGATTC	2700
Qy	2701	GGATCCT	CTTCTGTGGCCAGCGGCCCTTCGAGACCGCGGGGAGAACCCCTATGCTGT	2760

```
Db 2701 GGATCCTCTCTCTGTGGCCAGGCGCCCTCGAGCCCGGAGAACCCCTATGCTTGT 2760
Qy 2761 CCGGAATGGTAAAGTCTTCTCAGAGCTCTCAGCTGGTGGCCACAGCGGTACCCAC 2820
Db 2761 CCGGAATGGTAAAGTCTTCTCAGAGCTCTCAGCTGGTGGCCACAGCGGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTGGCAAACTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTGGCAAACTTTTAGCCAGTCCAGCTCC 2880
Qy 2881 CTGGTGGCCATCAAGGCACTATCTACTGGCGAGAGCCATACAAATGTCGAAATGTCG 2940
Db 2881 CTGGTGGCCATCAAGGCACTATCTACTGGCGAGAGCCATACAAATGTCGAAATGTCG 2940
Qy 2941 AAGTCTTCTCTCTCGGTCTGACAACTCGTTCGGCCACCAAGCTACTCACACCGGGAGAG 3000
Db 2941 AAGTCTTTCAGCAGCTCCAGCTCGGTTCGGCCACCAAGCTACTCACACCGGGAGAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGGTAAAGTCTTTCAGCCAGCGAGTAACTTGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGGTAAAGTCTTTCAGCCAGCGAGTAACTTGTGGCCAC 3060
Qy 3061 CAGCGTACCCACAGGCTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACAGGCTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGT 3120
Qy 3121 CAGCGCGGCCACTTGGCCAGCATCAAGCGCACTCATCTGCGAGAGCCATACAAATGT 3180
Db 3121 GACTCGCGGACCTTGTCCGCATCAAGCGCACTCATCTGCGAGAGCCATACAAATGT 3180
Qy 3181 CCAGAAATGGCAAGTCTTCTCTCGGTCTGACAACTCTGTCGGCCACCAACGTACTCAC 3240
Db 3181 CCAGAAATGGCAAGTCTTCTCTCCAAATCCAGCCATCTCGTCCGCGACCAACGTACTCAC 3240
Qy 3241 ACCGTAATAAAATAGTGGCCAGCGCGCCAGTACCGTACGAGCTTCGCGACTACGCT 3300
Db 3241 ACCGTAATAAAATAGTGGCCAGCGCGCCAGTACCGTACGAGCTTCGCGACTACGCT 3300

RESULT 6
AX284144 7373 bp DNA linear PAT 20-NOV-2001
LOCUS
DEFINITION Sequence 1 from Patent WO0178756.
ACCESSION AX284144
VERSION AX284144.1 GI:17044832
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1.
AUTHORS Wiedersanders,B. and Maubach,G.
TITLE Agent for postoperative use after the removal of bone tumours
JOURNAL Patent: WO 0178756-A 1 25-Oct-2001;
Depuy Biotech Jena GmbH (DE)
FEATURES
Location/Qualifiers
1..7373
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Expressionsvektor pMalc2cchbnp2"
1528..3453
/note="unnamed protein product; /produkt=Fusionsprotein
MBP Cystatin C BMP-2"
/codon_start=1
/transl_table=11
/protein_id="CAD12160.1"
/db_xref="GI:17044833"
/translation="MKIEBKLVINWKGNGVGLAEVGVKFKEDTGKIVVBPDKL
BEKFPQVATGGDDIIIFWADRFGGYAQSGLLAEITPDKAFQDKLYFPFVADRYNG
KLIAYPJAVBALSLYNKIDLLPNPKTWEEIPALDKELAKGKSLMFLNLDQSYFTWP
LIAADGGYAFKNGYKIDVGNAGAKAGLTLFVLDLTKNKMADTDYSIAEAA
NKGETAMTINGPMAWSNIDTSKVNYGVTVLPTFKGQSPKFPVGLSAGINAAAPNKEL
```

```
misc_feature
2689
/note="Faktor Xa Protease /position=388 (AMINO
ACIDsequenz)"
3067..3108
/misc_feature
/note="/spacer zwischen Cystatin C und BMP-2"
ORIGIN
Query Match 82.0%; Score 2705.8; DB 6; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGCGCCCGAAGAGA 60
Qy 61 GTCAAATTCAGGCTGGTGAATGTGAACCACTGTTATACGATGTGCGAGAGTATCCCG 120
Db 61 GTCAAATTCAGGCTGGTGAATGTGAACCACTGTTATACGATGTGCGAGAGTATCCCG 120
Qy 121 GTGTCCTTATCAGACCGTTTCCGGGTGTGAACCAAGCCAGCCAGCGTTCCTGCGAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCGGGTGTGAACCAAGCCAGCCAGCGTTCCTGCGAAA 180
Qy 181 CCGGGGAAAAAGTGGAGCGCGATGGCGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
Db 181 CCGGGGAAAAAGTGGAGCGCGATGGCGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGGAAACAGTGTGCTGATTCGGCTTGCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGGAAACAGTGTGCTGATTCGGCTTGCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACSCGCGCTCGCAAAATTTGTCGGCGGATTAATAATCTCGCGCGATCACTGGGTGCCAG 360
Db 301 ACSCGCGCTCGCAAAATTTGTCGGCGGATTAATAATCTCGCGCGATCACTGGGTGCCAG 360
Qy 361 TGGTGTGTGTCATGGTAGAACGAAGCGGCTCGAAGCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGTGTGTCATGGTAGAACGAAGCGGCTCGAAGCTGTAAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTCCTCGCTGGATGACAGAGTGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATTCCTCGCTGGATGACAGAGTGCCA 480
Qy 481 TTCTGTGGAGCTGCTGCACATAATTTTTCGGCGCTTATTTCTTGATGTCTTGACGAGA 540
Db 481 TTCTGTGGAGCTGCTGCACATAATTTTTCGGCGCTTATTTCTTGATGTCTTGACGAGA 540
Qy 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGACGGTACGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGACGGTACGGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db 601 TGGTGGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
Qy 661 CGGCTGTGGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Db 661 CGGCTGTGGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Qy 721 CGGAACGGGAAGCGGCTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATCTGA 780
Db 721 CGGAACGGGAAGCGGCTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTGCTGTTGCCAACGATCAGATGGCGCTGGCGGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTGCTGTTGCCAACGATCAGATGGCGCTGGCGGCAA 840
```

Qy	841	TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG	900
Db	841	TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACACACCATCAAAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACACACCATCAAAACAGGATTTTC	960
Qy	961	GCCTGCTGGGCAAAACGAGCTGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020
Db	961	GCCTGCTGGGCAAAACGAGCTGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020
Qy	1021	AGGGAATCAGCTGTTGCGGCTCTCACTGCTGGAAGAAAACCAACCCCTGGCGGCCAATA	1080
Db	1021	AGGGAATCAGCTGTTGCGGCTCTCACTGCTGGAAGAAAACCAACCCCTGGCGGCCAATA	1080
Qy	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Db	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Qy	1141	CCGACTCGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGACTCGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Qy	1201	GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCAAGGTAATGAGTTAGCTCACTCATTTAG	1260
Db	1201	GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCAAGGTAATGAGTTAGCTCACTCATTTAG	1260
Qy	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGTAATCACTGCATTAATTCG	1320
Db	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGTAATCACTGCATTAATTCG	1320
Qy	1321	TGTCCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCATACGGTT	1380
Db	1321	TGTCCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCATACGGTT	1380
Qy	1381	CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTTGA	1440
Db	1381	CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTTGA	1440
Qy	1441	ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCGGTTAGTGTTCACGA	1500
Db	1441	ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCGGTTAGTGTTCACGA	1500
Qy	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGGAAGAGTAAACTGGTAATCTGG	1560
Db	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGGAAGAGTAAACTGGTAATCTGG	1560
Qy	1561	ATTAAACGCGATAAAGGCTATAACGGTCTCGCTGAAAGTTCGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGCGATAAAGGCTATAACGGTCTCGCTGAAAGTTCGTAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGAAATTAAGTCAACGGTTGAGCATTCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680
Db	1621	ACCGAAATTAAGTCAACGGTTGAGCATTCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680
Qy	1681	GCGCAACTGGCGATGGCCCTGACATTTATCTTCGGCACACAGCCGCTTGTGGCTAC	1740
Db	1681	GCGCAACTGGCGATGGCCCTGACATTTATCTTCGGCACACAGCCGCTTGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
Db	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
Qy	1801	CCGTTTACCTGGATGCGGTAGCTTTACACGCAAGCTGATTGCTTACCGATCCCTGTT	1860
Db	1801	CCGTTTACCTGGATGCGGTAGCTTTACACGCAAGCTGATTGCTTACCGATCCCTGTT	1860
Qy	1861	GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
Qy	1921	GAGATCCGCGCTGGATTAAGAACTGAAACGGAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGCGCTGGATTAAGAACTGAAACGGAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTCCACCTGCGCTGATTCGCTGACGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCCACCTGCGCTGATTCGCTGACGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG	2100
Qy	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATGATGACAGACCGATTAC	2160
Db	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATGATGACAGACCGATTAC	2160
Qy	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Qy	2221	GCATGGTCCAAACATCGACACCAAGAGTGAATTAATGGTTAAACGGTACTGCGGACCTTC	2280
Db	2221	GCATGGTCCAAACATCGACACCAAGAGTGAATTAATGGTTAAACGGTACTGCGGACCTTC	2280
Qy	2281	AAGGTCMAACCATCAAAACCGTTGTTGGCGTGTGACGCGCAGGTATTAACGCCGCCAGT	2340
Db	2281	AAGGTCMAACCATCAAAACCGTTGTTGGCGTGTGACGCGCAGGTATTAACGCCGCCAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCCAAAACATATCTGCTGACGTGAAGGTCTG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCCAAAACATATCTGCTGACGTGAAGGTCTG	2400
Qy	2401	GAAGCGGTTAATAAGACAAACCCGCTGGGTCCGCTAGCGCTGAAGTCTTACGAGGAAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCCGCTGGGTCCGCTAGCGCTGAAGTCTTACGAGGAAG	2460
Qy	2461	TTGGGAAAGATCCACGTTATTCGCCGCACCATGGAACCGCCAGAAAGGTGAAATCATG	2520
Db	2461	TTGGGAAAGATCCACGTTATTCGCCGCACCATGGAACCGCCAGAAAGGTGAAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC	2580
Qy	2581	GCCAGCGTCTGACAGCTGCTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGTCTGACAGCTGCTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Qy	2641	AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGAGGATTTCAGAAATTC	2700
Db	2641	AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGAGGATTTCAGAAATTC	2700
Qy	2701	GGATCTCTTCTCTGTTGGCCGAGCGGCTTCGAGCCCGG 2741	
Db	2701	GGATCTCTTCTCTGTTGGCCGAGCGGCTTCGTTGGGAGG 2741	
RESULT 7			
AX172306			
LOCUS	AX172306	8101 bp	DNA linear PAT 03-JUL-2001
DEFINITION	Sequence 5 from Patent WO014467.		
ACCESSION	AX172306		
VERSION	AX172306.1	GI:14597487	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Goubin-Gramatica,F., Ducommun,B. and Prevost,G.		
TITLE	Method for obtaining human cdc25 phosphatases and method for		
JOURNAL	identifying human cdc25 phosphatase modulators		
	Patent: WO 014467-A 5 21-JUN-2001;		
	SOCIETE DE CONSEILS DE RECHERCHES ET D'APPLICATIONS SCIENTIFIQUES		

FEATURES		(FR)		Location/Qualifiers	
source		1..8101		/organism="synthetic construct"	
		/mol_type="unassigned DNA"			
		/db_xref="taxon:32630"			
ORIGIN		/note="bouche JM109 / pMAL - Hs Cdc25C"			
Query Match		82.0%; Score 2705.8; DB 6; Length 8101;			
Best Local Similarity		99.9%; Pred. No. 0;			
Matches 2707; Conservative		0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA	60		
DB	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA	60		
QY	61	GTCAATTGAGGTGGTGAATGTGAACACGATTAACGTTATACGATGTCGAGATGCGCG	120		
DB	61	GTCAATTGAGGTGGTGAATGTGAACACGATTAACGTTATACGATGTCGAGATGTCGCG	120		
QY	121	GTGTCCTTTATCAGACCGTTTTCCCGCGTGTGAACCGCCAGCCAGCTTTCTGCGAAAA	180		
DB	121	GTGTCCTTTATCAGACCGTTTTCCCGCGTGTGAACCGCCAGCCAGCTTTCTGCGAAAA	180		
QY	181	CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC	240		
DB	181	CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC	240		
QY	241	AACAACTGGCGGCAACAGTCGTGATTTGGCGTTGGCACTTCAGTCTGGCCCTGC	300		
DB	241	AACAACTGGCGGCAACAGTCGTGATTTGGCGTTGGCACTTCAGTCTGGCCCTGC	300		
QY	301	ACGCGCCGTTCGCAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGTGGCCAGC	360		
DB	301	ACGCGCCGTTCGCAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGTGGCCAGC	360		
QY	361	TGTTGGTTCGATGATAGAACGAGCGCGTGAAGCGCTGTAAGCGCGGTGCAACAATC	420		
DB	361	TGTTGGTTCGATGATAGAACGAGCGCGTGAAGCGCTGTAAGCGCGGTGCAACAATC	420		
QY	421	TTCTCGCGCAACGGTTCAGTGGCGTATTAATTAATTAATTAATTAATTAATTAATTA	480		
DB	421	TTCTCGCGCAACGGTTCAGTGGCGTATTAATTAATTAATTAATTAATTAATTAATTA	480		
QY	481	TTGCTGTGGAAGTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGAC	540		
DB	481	TTGCTGTGGAAGTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGAC	540		
QY	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600		
DB	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600		
QY	601	TGTTGCGATTCGGTTCACGCAAAATCGCGTGTAGCGGCGCCATTAAGTCTGTCTCGG	660		
DB	601	TGTTGCGATTCGGTTCACGCAAAATCGCGTGTAGCGGCGCCATTAAGTCTGTCTCGG	660		
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
QY	721	CGGAACGGGAAGGCACTGGAGTGGCATGTCGGTTTTTCAACAAACCAATGCAATGCTGA	780		
DB	721	CGGAACGGGAAGGCACTGGAGTGGCATGTCGGTTTTTCAACAAACCAATGCAATGCTGA	780		
QY	781	ATGAGGCGATCGTTCCCACTCGCATGCTGTTGCAACGATTCAGTGGCGTGGCGCA	840		
DB	781	ATGAGGCGATCGTTCCCACTCGCATGCTGTTGCAACGATTCAGTGGCGTGGCGCA	840		
QY	841	TGCGCGCATTCAGATTCGCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900		
DB	841	TGCGCGCATTCAGATTCGCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900		
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAAAACAGGATTTTC	960		
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAAAACAGGATTTTC	960		
QY	961	GCCTGCTGGGGCAAAACAGCGTGCACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA	1020		
DB	961	GCCTGCTGGGGCAAAACAGCGTGCACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA	1020		
QY	1021	AGGCAATTCAGCTGTTGCCGTCTCACTGTGTAAGAAAGAAACACCTTGGCGCCCAATA	1080		
DB	1021	AGGCAATTCAGCTGTTGCCGTCTCACTGTGTAAGAAAGAAACACCTTGGCGCCCAATA	1080		
QY	1081	CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCAGCTGGCAGCAGGTTT	1140		
DB	1081	CGCAACCGCTCTCTCCCGCGGCTTGGCGGATTCATTAATGCAGCTGGCAGCAGGTTT	1140		
QY	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTATTAG	1200		
DB	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTATTAG	1200		
QY	1201	GCACAATTCATGTTTGAACAGCTTATCATCGACTGCAACGCGTGCAAAATGCTTGGCG	1260		
DB	1201	GCACAATTCATGTTTGAACAGCTTATCATCGACTGCAACGCGTGCAAAATGCTTGGCG	1260		
QY	1261	TCAGGAGCCATCGAAAGCTGTGTATGGCTGTGCAAGTCTGTAATCACTGCAATAATTCG	1320		
DB	1261	TCAGGAGCCATCGAAAGCTGTGTATGGCTGTGCAAGTCTGTAATCACTGCAATAATTCG	1320		
QY	1321	TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCCGACATCAAAACGGTT	1380		
DB	1321	TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCCGACATCAAAACGGTT	1380		
QY	1381	CTGGCAATATCTGAAATGAGCTGTGTGCAATTAATTAATTAATTAATTAATTAATTA	1440		
DB	1381	CTGGCAATATCTGAAATGAGCTGTGTGCAATTAATTAATTAATTAATTAATTAATTA	1440		
QY	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGGTTTTCACGA	1500		
DB	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGGTTTTCACGA	1500		
QY	1501	GCACTTCAACAAAGGACCATAGATTATGAAATCGAAGAGGTAACTGGTAAATCTCGG	1560		
DB	1501	GCACTTCAACAAAGGACCATAGATTATGAAATCGAAGAGGTAACTGGTAAATCTCGG	1560		
QY	1561	ATTAAAGCGGATTAAGGCTATTAACGCTCTCGCTGAAAGTAAAGTAAAGTAAAGTAA	1620		
DB	1561	ATTAAAGCGGATTAAGGCTATTAACGCTCTCGCTGAAAGTAAAGTAAAGTAAAGTAA	1620		
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAAGTAAAGTAAAGTAAAGTAAAGT	1680		
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAAGTAAAGTAAAGTAAAGTAAAGT	1680		
QY	1681	CGCGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACGACCGCTTTGGTGGCTAC	1740		
DB	1681	CGCGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACGACCGCTTTGGTGGCTAC	1740		
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGCAAAAGCGTTCCAGGACAGCTGAT	1800		
DB	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGCAAAAGCGTTCCAGGACAGCTGAT	1800		
QY	1801	CCGTTTACCTGGGATCCCGTACGTTACAAACGCAAGTGTATGCTTACCCGATCGCTGTT	1860		
DB	1801	CCGTTTACCTGGGATCCCGTACGTTACAAACGCAAGTGTATGCTTACCCGATCGCTGTT	1860		
QY	1861	GAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCAAAACCTGGGAA	1920		
DB	1861	GAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCAAAACCTGGGAA	1920		
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980		
DB	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980		
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG	2040		

Db 1981 |||||CTGCAGAAACCGTACTTTCACCTGGCCGCTGATTGCTGTGACGGGGGTTATCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAACAACAACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAACAACAACATGAATGCGAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACACGCAAGTGAATTAATGGTGAACGCTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGCAAGTGAATTAATGGTGAACGCTACTGCGGACCTTC 2280
Qy 2281 AAGGTCAACCTCAACCGTTCTTGGCGTGTGAGGCGAGGTATTAAACGCCCGCAGT 2340
Db 2281 AAGGTCAACCTCAACCGTTCTTGGCGTGTGAGGCGAGGTATTAAACGCCCGCAGT 2340
Qy 2341 CGGAACAAAGAGCTGCGAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CGGAACAAAGAGCTGCGAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGTCCACGATTATTCGCCGCCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGTCCACGATTATTCGCCGCCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CGGAACATCCCGCAGATGTCGCTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CGGAACATCCCGCAGATGTCGCTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAAGACGCCGACAGTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAAGACGCCGACAGTAAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
AX377531
LOCUS AX377531 6648 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 8 from Patent WO0212553.
ACCESSION AX377531
VERSION AX377531.1 GI:19573717
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensorf,H. and Muth,J.
TITLE Method for detecting mutations in nucleotide sequences
JOURNAL Patent: WO 0212553-A 8 14-FEB-2002;
Nanogen Recognomics GmbH (DE)
FEATURES
1. 6648
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN /note="Plasmid pMAL-c2x"
Query Match 81.8%; Score 2701; DB 6; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGCGCATGATAGGCCCGCGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGCGCATGATAGGCCCGCGAAGAGA 60
Qy 61 GTCAATTCCAGGTGTGAATGTGAACCCAGTAAAGTATACGATGTCGACAGATATGCG 120
Db 61 GTCAATTCCAGGTGTGAATGTGAACCCAGTAAAGTATACGATGTCGACAGATATGCG 120
Qy 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGTAACACGAGCCAGCCACGTTTCTGCGAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGTAACACGAGCCAGCCACGTTTCTGCGAAAA 180
Qy 181 CGCGGAAAAAGTGGAGCGCGGATGGCGGAGCTGAATTATATCCCAACCGCTGGCAC 240
Db 181 CGCGGAAAAAGTGGAGCGCGGATGGCGGAGCTGAATTATATCCCAACCGCTGGCAC 240
Qy 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCCCTCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCCTCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGGTGTGATGATGAAGACGAGCGCGCTGGAAGCCTGTAAGCGCGGTGCAATATC 420
Db 361 TGGTGGTGTGATGATGAAGACGAGCGCGCTGGAAGCCTGTAAGCGCGGTGCAATATC 420
Qy 421 TTCTCGCGCAACGCTGATGCGGCTGATCAATTAATCTCCGCTGATGATGATGATGATGATGAT 480
Db 421 TTCTCGCGCAACGCTGATGCGGCTGATCAATTAATCTCCGCTGATGATGATGATGATGATGAT 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTAATTTTCTCCATGAAGACGCTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTTCTCCATGAAGACGCTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGTCCACGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGCTCGG 660
Db 601 TGGTGGCATTTGGTCCACGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGCTCGG 660
Qy 661 CGCGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 720
Qy 721 CGGAACGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTCGGATGCTGTTGCAACGATCAGATGCGGTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTCGGATGCTGTTGCAACGATCAGATGCGGTGGCGGCA 840
Qy 841 TGCGGCCATTACCGAGTCCCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TGCGGCCATTACCGAGTCCCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGTGCTCAACTCTCTCAGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGTGCTCAACTCTCTCAGGCCAGCGGTGA 1020

```
Db 961 GCCTGCTGGGCAAAACGAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGTCCTGCTCACTGCTGAAAGAAACCAACCTCGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTCCTGCTCACTGCTGAAAGAAACCAACCTCGGCGCCCAATA 1080
Qy 1081 CGGAACCGCTCTCTCCCGCGGTGGCGATTCATTAATGACGTGGCAGCAGAGTTT 1140
Db 1081 CGGAACCGCTCTCTCCCGCGGTGGCGATTCATTAATGACGTGGCAGCAGAGTTT 1140
Qy 1141 CCCGCTGAAAGCGGAGTGCAGCGCAACGCAATTAATGAGTGTAGCTACTCATTTAG 1200
Db 1141 CCCGCTGAAAGCGGAGTGCAGCGCAACGCAATTAATGAGTGTAGCTACTCATTTAG 1200
Qy 1201 GCACAATCTCATTTGACAGCTTATCATCACTGACGCGTGCACCAATGCTTCGCG 1260
Db 1201 GCACAATCTCATTTGACAGCTTATCATCACTGACGCGTGCACCAATGCTTCGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGGTATGCTGTGAGGTCTGTAATCACTGCAATTTTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGGTATGCTGTGAGGTCTGTAATCACTGCAATTTTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTGATATGTTTTCGCGGCAATCATACGTTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGATATGTTTTCGCGGCAATCATACGTTT 1380
Qy 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATTCG 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATTCG 1440
Qy 1441 ATTGTAGCGGATACAAATTTACACAGGAAACGCCAGTCTGTTAGGTGTTTCAG 1500
Db 1441 ATTGTAGCGGATACAAATTTACACAGGAAACGCCAGTCTGTTAGGTGTTTCAG 1500
Qy 1501 GCACCTTCCAAACAGGACCATAGATTATGAAACTGGAAGAGTAAACTCGTAACTCG 1560
Db 1501 GCACCTTCCAAACAGGACCATAGATTATGAAACTGGAAGAGTAAACTCGTAACTCG 1560
Qy 1561 ATTAACGGCGATAAAGGTATAACCGTCTCGCTGAACTCGGTGAGAAATTCGAGAA 1620
Db 1561 ATTAACGGCGATAAAGGTATAACCGTCTCGCTGAACTCGGTGAGAAATTCGAGAA 1620
Qy 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGACAG 1680
Db 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGACAG 1680
Qy 1681 GCGCAACTGCGGATGGCCCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGGCCCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGGACAAAGGTTCCAGGACAACTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGGACAAAGGTTCCAGGACAACTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACCGCAAGCTGATTGCTACCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACCGCAAGCTGATTGCTTACCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGTCGCGAACCGCCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGTCGCGAACCGCCCAAAACCTCGGAA 1920
Qy 1921 GAGATCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGCGCGCTGATGTTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTCCTGCTGACGGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTCCTGCTGACGGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGGCGTGAATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGGCGTGAATACGCTGGCGGAAAGCG 2100
```

```
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGGAAACAGCATGACCATCAACGGCCCGGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGGAAACAGCATGACCATCAACGGCCCGGTGG 2220
Qy 2221 GCATGTGTCACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCGCACCTTC 2280
Db 2221 GCATGTGTCACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCGCACCTTC 2280
Qy 2281 AAGGCTCAACCATCAACACCGTTTGGTGGCGTGTGAGCGCAGGTATTAACGGCCCGCAGT 2340
Db 2281 AAGGCTCAACCATCAACACCGTTTGGTGGCGTGTGAGCGCAGGTATTAACGGCCCGCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGTTAATAAAGCAAAACCGCTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTAATAAAGCAAAACCGCTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATTCGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTCGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGTGTCCGCTTTCTGGTATGCGCTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGTGTCCGCTTTCTGGTATGCGCTGCTGCTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTGTGAGACTGTCGATGAAGCCCTGAAAGACCGCAGACACTAATTCGAGTCG 2640
Db 2581 GCCAGCGTGTGAGACTGTCGATGAAGCCCTGAAAGACCGCAGACACTAATTCGAGTCG 2640
Qy 2641 AACAAACAAACAAATAACAAATAACAACTCGGATCGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAACAAATAACAACTCGGATCGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCTCTT 2709
Db 2701 GGATCTCTT 2709

RESULT 9
LOCUS A91965 7475 bp DNA circular PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9822496.
ACCESSION A91965
VERSION A91965.1 GI:6740811
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood,M.R. and Hurst,D.N.
TITLE ANTIVIRAL PEPTIDE DERIVATIVES
JOURNAL Patent: WO 9822496-A 1 28-MAY-1998;
HOFFMANN LA ROCHE (CH)
FEATURES
source Location/Qualifiers
1..7475
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATCGTGTCAAAACCTTTTCGCGGTATGCGATGATAGCGCCCGGAAGAGA 60
```

Db 1 CCGACACCATGGAATGGTGGCAAAACCTTTTCGCGGTATGGCATGATGAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCCGAGAGTATGCCG 120
Db 61 GTCAATTTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCCGAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Qy 241 AACAACTCGCGGCAAAACAGTGTGCTGATTTGGCGTTTGCACCTTCCAAGTCTGGGCCCTGC 300
Db 241 AACAACTCGCGGCAAAACAGTGTGCTGATTTGGCGTTTGCACCTTCCAAGTCTGGGCCCTGC 300
Qy 301 ACGGCGCGTCGCAAAATTGTTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCGCAGCG 360
Db 301 ACGGCGCGTCGCAAAATTGTTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCGCAGCG 360
Qy 361 TGGTGTGTCGATGTGAAGCGAGCGCGTGAAGCGCTGTAAAGCGCGGTGCAATC 420
Db 361 TGGTGTGTCGATGTGAAGCGAGCGCGTGAAGCGCTGTAAAGCGCGGTGCAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATCACCAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATCACCAGGATGCCA 480
Qy 481 TTGTGTGGAAGCTGCCTGCACTAAATGTTTCGGCGGTTATTTCTTGATGTCTGACACAGA 540
Db 481 TTGTGTGGAAGCTGCCTGCACTAAATGTTTCGGCGGTTATTTCTTGATGTCTGACACAGA 540
Qy 541 CACCCATCAACGATTAATTTCTCCCATGAAGACGGTACGGCATGGCGGTGGAGCATC 600
Db 541 CACCCATCAACGATTAATTTCTCCCATGAAGACGGTACGGCATGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGTCAACAGAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGCATTTGGTCAACAGAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGACCTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGCGACCTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCGCATGGCGTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCGCATGGCGTGGGCGCAA 840
Qy 841 TCGCGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAGC 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAGC 900
Qy 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGGTTAAACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGGTTAAACCATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGACGTGGCAACAGAGTTT 1140

Db 1081 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGACAGTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGCTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGCTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACAGCCATCGGAAGCTGTGATGGCTGTGAGGTCTGAGTCTGAGTCACTGCAATAATTCG 1320
Db 1261 TCAGGACAGCCATCGGAAGCTGTGATGGCTGTGAGGTCTGAGTCTGAGTCACTGCAATAATTCG 1320
Qy 1321 TGTCTGCTCAAGCGCGACATCCCGTTCTGGAATATGTTTTTTCGCGCGACATCATTAACGGTT 1380
Db 1321 TGTCTGCTCAAGCGCGACATCCCGTTCTGGAATATGTTTTTTCGCGCGACATCATTAACGGTT 1380
Qy 1381 CTGGCAAAATCTCTGAAATGAGCTTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAAAATCTCTGAAATGAGCTTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCACGA 1500
Qy 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 CGGCAACTGCGCATGGCCCTGACATTAATCTTCTGGGCACACGCCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGCATGGCCCTGACATTAATCTTCTGGGCACACGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGACAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTCGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTCGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAAACCCTGGGAA 1920
Db 1861 GAAGCGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAAACCCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCCACTGCGCGCTGATTTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACTGCGCGCTGATTTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGTTGACCTGATTTAAAAACAAACATGAATGCGACACCGATTTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTTAAAAACAAACATGAATGCGACACCGATTTAC 2160
Qy 2161 TCCATCGCAGAGCTGCGCTTTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220

QY 1321 TGTGCTCAAGGGGCACTCCCGTTCTGGATAATGTTTTTGGCCGCAATCATATAAGGTT 1380
Db 1321 TGTGCTCAAGGGGCACTCCCGTTCTGGATAATGTTTTTGGCCGCAATCATATAAGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATATTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCGCTTTAGGTGTTTTTCAACA 1500
Db 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCGCTTTAGGTGTTTTTCAACA 1500
QY 1501 GCACCTCACCAACAGGACCATAGATTATGAAAACTGAAAGAGGTAAACTGGTATATCTGG 1560
Db 1501 GCACCTCACCAACAGGACCATAGATTATGAAAACTGAAAGAGGTAAACTGGTATATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTTAAAGTTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTTAAAGTTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCAGGTT 1680
QY 1681 GGGGCACTGGCGATGGCCCTGACATTATCTCTGGGCAACGACCGCTTTGGTGGCTTAC 1740
Db 1681 GGGGCACTGGCGATGGCCCTGACATTATCTCTGGGCAACGACCGCTTTGGTGGCTTAC 1740
QY 1741 GGTCAATCTGGCTGTTGGCTGAAATCACCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GGTCAATCTGGCTGTTGGCTGAAATCACCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTGCTTTACCCGATCGCTGT 1860
Db 1801 CGGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTGCTTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGTGTATTTATACAAAGATCTGTCGCGAAACCGGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGTGTATTTATACAAAGATCTGTCGCGAAACCGGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATAAAGACTGAAGCGAAGGTAAAGCGCGTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAAGACTGAAGCGAAGGTAAAGCGCGTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACACATTAAGACGTTGGCGTGGAATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACACATTAAGACGTTGGCGTGGAATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGATGACACACCGATTAC 2160
QY 2161 TCCATCGCAAGAGCTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCGCTGG 2220
Db 2161 TCCATCGCAAGAGCTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACCGAAAGTGAATTTATGTTGTTAAACGCTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCGAAAGTGAATTTATGTTGTTAAACGCTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCATCCAAACCGTTGCTGGCTGCTGAGCGCAGGATTTAAACGCGCCAGT 2340
Db 2281 AAGGTCACCATCCAAACCGTTGCTGGCTGCTGAGCGCAGGATTTAAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTCATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTCATGAAGTCTG 2400

QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTAGCGCTGAGCTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTAGCGCTGAGCTCTTACGAGGAAGAG 2460
QY 2461 TTGGGAAAAGATCCACGCTATTGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGAAAAGATCCACGCTATTGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAAACAATAACAACAACTCGGGATCGAGGGAAAGATTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAACAATAACAACAACTCGGGATCGAGGGAAAGATTTCAGAAATTC 2700

RESULT 11
AR207294
LOCUS AR207294 7475 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6372883.
ACCESSION AR207294
VERSION AR207294.1 GI:21506162
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7475)
AUTHORS Atwood,M,Richard., Hurst,D,Nigel., Jones,P,Stephen.,
Kay,P,Brittain., Raynham,T,Michael. and Wilson,F,Xavier.
TITLE Antiviral medicaments
JOURNAL Patent: US 6372883-A 1 16-APR-2002;
FEATURES
source Location/Qualifiers
1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCACACCATCGAATGGTGCAAAACCTTTCCGGGTATGCGATAGCGCCCGGAAGAGA 60
Db 1 CGCACACCATCGAATGGTGCAAAACCTTTCCGGGTATGCGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTGAGGTGTTGAATGTGAAACCAAGTAACTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTGAGGTGTTGAATGTGAAACCAAGTAACTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCCTTTATCAGACCGTTTCCCGCGTGTGTAACCAAGCCAGCCACGCTTTCTGCGAAAA 180
Db 121 GTGTCCTTTATCAGACCGTTTCCCGCGTGTGTAACCAAGCCAGCCACGCTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTCAATTTACATTCACACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTCAATTTACATTCACACCGCGTGGCAC 240
QY 241 AACAACTGCGGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 AGCGCCGCTCGCAAAATTTGTCGGCGGATTAATCTCGGCCGATCAACTGGGTGCGCAGCG 360
Db 301 AGCGCCGCTCGCAAAATTTGTCGGCGGATTAATCTCGGCCGATCAACTGGGTGCGCAGCG 360
QY 361 TGGTGGTCTGATGCTAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGGTGCACATTC 420
Db 361 TGGTGGTCTGATGCTAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGGTGCACATTC 420

Qy	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACACAGGATGCCA	480
Db	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACACAGGATGCCA	480
Qy	481	TTGCTGTGGAAAGCTCGCTGCGACTAATAGTTCGGCGTTATTCTTGTATGTCCTGACCCAGA	540
Db	481	TTGCTGTGGAAAGCTCGCTGCGACTAATAGTTCGGCGTTATTCTTGTATGTCCTGACCCAGA	540
Qy	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGGCACTGGGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGGCACTGGGCGTGGAGCATC	600
Qy	601	TGGTCGCAATTGGGTACACAGCAAAATCGCGCTGTTTAGCGGGCCCAATAAGTTCGTGCTCGG	660
Db	601	TGGTCGCAATTGGGTACACAGCAAAATCGCGCTGTTTAGCGGGCCCAATAAGTTCGTGCTCGG	660
Qy	661	CGCGTCTCGGCTCGCTGGCTGGCATAAATATCTCACTCGCMAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTCGGCTCGCTGGCTGGCATAAATATCTCACTCGCMAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAAACGGGAAGGCGACTGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAAACGGGAAGGCGACTGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780
Qy	781	ATGAGGGCATCGTTCCCACTCGGATCTGGTTGCCAAGATCAGATGGCGCTGGGCGCAA	840
Db	781	ATGAGGGCATCGTTCCCACTCGGATCTGGTTGCCAAGATCAGATGGCGCTGGGCGCAA	840
Qy	841	TGCGCGCATTTACCGAGTCCGGGCTGGCGTTGGTGCGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACCGAGTCCGGGCTGGCGTTGGTGCGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACGGAAGACAGTCTATGTTATATCCGCGGTTAAACCAATCAAAACAGGATTTTC	960
Db	901	ACGATACGGAAGACAGTCTATGTTATATCCGCGGTTAAACCAATCAAAACAGGATTTTC	960
Qy	961	GCCTGCTGGGGCAAAACAGCGTGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAACAGCGTGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
Qy	1021	AGGCAATACGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTCGCGCCCAATA	1080
Db	1021	AGGCAATACGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTCGCGCCCAATA	1080
Qy	1081	CGAAACCGGCTCTCCCGCGGTTGGCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Db	1081	CGAAACCGGCTCTCCCGCGGTTGGCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Qy	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Qy	1201	GCACAATCTCATGTTTGACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGGCG	1260
Db	1201	GCACAATCTCATGTTTGACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGGCG	1260
Qy	1261	TCAGGCAGCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATACCTGCAATAATTCG	1320
Db	1261	TCAGGCAGCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATACCTGCAATAATTCG	1320
Qy	1321	TGTCGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTTTGGCGCGACATCATACCGGTT	1380
Db	1321	TGTCGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTTTGGCGCGACATCATACCGGTT	1380
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTAAATGTTGGA	1440
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTAAATGTTGGA	1440
Qy	1441	ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTACCGA	1500
Db	1441	ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTACCGA	1500

QY	1501	GCAC	TTCA	CCAA	CA	AGG	ACC	CA	TAG	ATT	ATG	AAAA	CT	GA	GA	AG	GTAA	CT	GG	TAAT	CT	GG	1560				
Db	1501	GCAC	TTCA	CCAA	CA	AGG	ACC	CA	TAG	ATT	ATG	AAAA	CT	GA	GA	AG	GTAA	CT	GG	TAAT	CT	GG	1560				
QY	1561	ATT	AA	CGG	CG	ATA	AGG	CT	ATA	CGG	CT	CG	CT	GAA	GT	CGG	TAA	GAA	AA	TT	CG	AG	AA	AG	AT	1620	
Db	1561	ATT	AA	CGG	CG	ATA	AGG	CT	ATA	CGG	CT	CG	CT	GAA	GT	CGG	TAA	GAA	AA	TT	CG	AG	AA	AG	AT	1620	
QY	1621	ACC	GAA	TT	AA	AG	T	CAC	CG	TT	GAG	CAT	CCG	GAT	AAA	CT	GG	AA	GAGA	AA	TT	CC	CA	CAG	GT	1680	
Db	1621	ACC	GAA	TT	AA	AG	T	CAC	CG	TT	GAG	CAT	CCG	GAT	AAA	CT	GG	AA	GAGA	AA	TT	CC	CA	CAG	GT	1680	
QY	1681	GGG	CAA	CT	G	CG	AT	G	CG	CT	G	CA	T	AT	CT	T	GG	G	CA	CAG	C	CG	CT	T	GG	T	1740
Db	1681	GGG	CAA	CT	G	CG	AT	G	CG	CT	G	CA	T	AT	CT	T	GG	G	CA	CAG	C	CG	CT	T	GG	T	1740
QY	1741	GCT	CA	AT	CT	G	G	C	T	T	G	G	C	T	G	AA	AT	CA	C	C	C	G	G	A	C	G	1800
Db	1741	GCT	CA	AT	CT	G	G	C	T	T	G	G	C	T	G	AA	AT	CA	C	C	C	G	G	A	C	G	1800
QY	1801	CCG	TT	T	AC	T	CG	G	AT	G	CG	CA	AG	CT	G	AT	T	CA	A	CG	CA	AG	CT	G	AT	T	1860
Db	1801	CCG	TT	T	AC	T	CG	G	AT	G	CG	CA	AG	CT	G	AT	T	CA	A	CG	CA	AG	CT	G	AT	T	1860
QY	1861	GAA	G	CT	T	AT	CG	C	T	G	AT	T	TA	A	CA	AA	GA	T	C	G	T	C	CA	A	C	C	1920
Db	1861	GAA	G	CT	T	AT	CG	C	T	G	AT	T	TA	A	CA	AA	GA	T	C	G	T	C	CA	A	C	C	1920
QY	1921	GAG	AT	C	C	G	G	C	G	CT	G	AT	A	A	GA	CT	G	A	A	A	G	G	T	A	A	G	1980
Db	1921	GAG	AT	C	C	G	G	C	G	CT	G	AT	A	A	GA	CT	G	A	A	A	G	G	T	A	A	G	1980
QY	1981	CTG	CA	GA	AA	CA	CG	T	ACT	T	C	A	C	T	G	C	G	C	G	C	T	G	A	T	T	G	2040
Db	1981	CTG	CA	GA	AA	CA	CG	T	ACT	T	C	A	C	T	G	C	G	C	G	C	T	G	A	T	T	G	2040
QY	2041	TAT	GA	AA	CA	CG	CA	AG	T	AC	G	A	CT	T	AA	AG	CA	GT	G	G	G	G	T	G	A	A	2100
Db	2041	TAT	GA	AA	CA	CG	CA	AG	T	AC	G	A	CT	T	AA	AG	CA	GT	G	G	G	G	T	G	A	A	2100
QY	2101	G	G	T	C	T	G	A	C	T	T	G	A	C	T	T	TA	A	A	A	A	A	A	A	A	A	2160
Db	2101	G	G	T	C	T	G	A	C	T	T	G	A	C	T	T	TA	A	A	A	A	A	A	A	A	A	2160
QY	2161	TCC	AT	CG	CA	GA	AG	CT	G	C	T	T	TA	T	AA	AG	G	CA	AA	C	A	G	C	A	T	A	2220
Db	2161	TCC	AT	CG	CA	GA	AG	CT	G	C	T	T	TA	T	AA	AG	G	CA	AA	C	A	G	C	A	T	A	2220
QY	2221	G	C	A	T	G	T	T	C	C	A	C	A	G	A	A	G	T	A	T	A	T	A	T	A		
Db	2221	G	C	A	T	G	T	T	C	C	A	C	A	G	A	A	G	T	A	T	A	T	A	T	A		
QY	2281	A	A	G	G	T	C	A	A	C	G	T	T	TA	T	AA	AG	G	CA	AA	C	A	G	C	A	T	2340
Db	2281	A	A	G	G	T	C	A	A	C	G	T	T	TA	T	AA	AG	G	CA	AA	C	A	G	C	A	T	2340
QY	2341	CCG	AA	CA	AA	G	AT	G	CG	CT	G	AT	T	CG	AA	A	A	A	A	A	A	A	A	A	A	A	2400
Db	2341	CCG	AA	CA	AA	G	AT	G	CG	CT	G	AT	T	CG	AA	A	A	A	A	A	A	A	A	A	A	A	2400
QY	2401	G	A	A	G	G	T	T	A	A	A	G	A	C	A	A	A	C	G	C	T	G	G	G	T	A	2460
Db	2401	G	A	A	G	G	T	T	A	A	A	G	A	C	A	A	A	C	G	C	T	G	G	G	T	A	2460
QY	2461	TTG	CG	AA	A	G	A	T	C	C	A	A	C	CA	T	G	G	A	A	A	C	G	C	C	A	A	2520
Db	2461	TTG	CG	AA	A	G	A	T	C	C	A	A	C	CA	T	G	G	A	A	A	C	G	C	C	A	A	2520
QY	2521	CCG	AA	C	A	T	CCG	C	A	G	A	T	G	T	T	C	G	A	A	A	A	A	A	A	A	A	2580
Db	2521	CCG	AA	C	A	T	CCG	C	A	G	A	T	G	T	T	C	G	A	A	A	A	A	A	A	A	A	2580
QY	2581	G	C	A	G	C	G	G	T	C	G	A	C	T	G	A	T	G	A	A	G	C	C	T	A	A	2640

Dn	2591	GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGACGGCGACACTAAATTCGAGCTCG	2640	Dn	541	CACCCATCAACAGTATTATTTTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
Qy	2641	AACAAACAACAATAACAATAACAACCTCGGGATCGAGGAGGAGGATTTGAGAAATTC	2700	Qy	601	TGTGCGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG	660
Dn	2641	AACAAACAACAATAACAATAACAACCTCGGGATCGAGGAGGAGGATTTGAGAAATTC	2700	Dn	601	TGTGCGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG	660
RESULT 12				Qy	661	CGCGTCTGCGTCTGGCTGGCGATATAATCTCACTCGCAATCAAAATTCAGCGGATAG	720
AX377532				Dn	661	CGCGTCTGCGTCTGGCTGGCGATATAATCTCACTCGCAATCAAAATTCAGCGGATAG	720
DEFINITION	AX377532	Sequence 9 from Patent WO0212553.		Qy	721	CGGAACGGGAAGGAGCTGGAGTGCATGTCGGGTTTTCAACAACCAACATGCAATTCGTGA	780
ACCESSION	AX377532			Dn	721	CGGAACGGGAAGGAGCTGGAGTGCATGTCGGGTTTTCAACAACCAACATGCAATTCGTGA	780
VERSION	AX377532.1	GI:19573718		Qy	781	ATGAGGCGCATGTTCCCACTGCGATGCTGTTGCAACGATCAGATGCGCTGGGCGCAA	840
KEYWORDS		synthetic construct		Dn	781	ATGAGGCGCATGTTCCCACTGCGATGCTGTTGCAACGATCAGATGCGCTGGGCGCAA	840
SOURCE		synthetic construct		Qy	841	TGCGGCGCATTTACGAGTTCGCGGCTGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATACG	900
ORGANISM		other sequences; artificial sequences.		Dn	841	TGCGGCGCATTTACGAGTTCGCGGCTGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATACG	900
REFERENCE				Qy	901	ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC	960
AUTHORS	Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdoerf,H. and Muth,J.			Dn	901	ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC	960
TITLE	Method for detecting mutations in nucleotide sequences			Qy	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
JOURNAL	Patent: WO 0212553-A 9 14-FEB-2002;			Dn	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
FEATURES	Nanogen Recognomics GmbH (DE)			Qy	1021	AGGGCAATCAGCTGTGTGCCCTCTCACTGTGTAAGAAAAACCAACCTGGCGCCCAATA	1080
source	Location/Qualifiers			Dn	1021	AGGGCAATCAGCTGTGTGCCCTCTCACTGTGTAAGAAAAACCAACCTGGCGCCCAATA	1080
	1..9191	/organism="synthetic construct"		Qy	1081	CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGTGCACGACAGGTTT	1140
		/mol_type="unassigned DNA"		Dn	1081	CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGTGCACGACAGGTTT	1140
		/db_xref="taxon:32630"		Qy	1141	CCGACTTGAAGAGGCGGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
		/note="pMALc2x-mutS Plaemid"		Dn	1141	CCGACTTGAAGAGGCGGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
ORIGIN				Qy	1201	GCACAAATTCATGTTTGACAGCTTATCATCGATGCAACGCGTGCAACCAATGCTTCGCGG	1260
	Query Match	81.8%;	Score 2698;	Dn	1201	GCACAAATTCATGTTTGACAGCTTATCATCGATGCAACGCGTGCAACCAATGCTTCGCGG	1260
	Best Local Similarity	99.8%;	Pred. No. 0;	Qy	1261	TCAGGACCCATCGGAGCTGTGGTATGCGTGTGCGAGTCTGAATCACTGCATTAATTCG	1320
	Matches 2701;	Conservative	0; Mismatches	Dn	1261	TCAGGACCCATCGGAGCTGTGGTATGCGTGTGCGAGTCTGAATCACTGCATTAATTCG	1320
			5; Indels	Qy	1321	TGTGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCATACCGTT	1380
			0; Gaps	Dn	1321	TGTGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCATACCGTT	1380
				Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTTG	1440
				Dn	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTTG	1440
				Qy	1441	ATTGTGAGCGGATTAACATTTTCAACAGGAAACAGCCAGTCCGTTAGGTTGTTTTCAGGA	1500
				Dn	1441	ATTGTGAGCGGATTAACATTTTCAACAGGAAACAGCCAGTCCGTTAGGTTGTTTTCAGGA	1500
				Qy	1501	GCACCTTCAACCAAGGACCATAGATATGAAAACCTGAAGAGGTAATACTGTAATCTTCG	1560
				Dn	1501	GCACCTTCAACCAAGGACCATAGATATGAAAACCTGAAGAGGTAATACTGTAATCTTCG	1560
				Qy	1561	ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620
				Dn	1561	ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620
				Qy	1621	ACCGGAATTAAGTACCCTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
				Dn	1621	ACCGGAATTAAGTACCCTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680

Db 1621 ACCGAAATTAAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCACACAGGTT 1680
Qy 1681 GCGCAACTGCGGATGCCCTGACATTATCTCTGGGCACACACCGCTTTGTTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGCCCTGACATTATCTCTGGGCACACACCGCTTTGTTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGGAATCACCCCGGACAAAGCGTTCCAGGACAACTGTTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGGAATCACCCCGGACAAAGCGTTCCAGGACAACTGTTAT 1800
Qy 1801 CCGTTTACCTGGGATGCCGTACGTTACAGCGGAAGCTGATTCCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCCGTACGTTACAGCGGAAGCTGATTCCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGTGATTTATAACAAGATCTCTCCGAACCCGCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGTGATTTATAACAAGATCTCTCCGAACCCGCCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGCTGACGGGGTTATCGCTTCAAG 2040
Db 1981 CTCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGCTGACGGGGTTATCGCTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGACCTGATTAATAAACAACATGAATGATGACGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGACCTGATTAATAAACAACATGAATGATGACGACACCGATTAC 2160
Qy 2161 TCCATCGCAAGAGCTGCCCTTAAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAAGAGCTGCCCTTAAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Qy 2221 GCATGCTCAACATCGACACCGCAAGAGTAAATATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCGCAAGAGTAAATATGTTGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAAAACCGTTCTGCGCTGCTGAGGCGAGGTATTAAAGCGCCGCTG 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTCTGCGCTGCTGAGGCGAGGTATTAAAGCGCCGCTG 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGCTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGCTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCAAGATCCACGTTATTCGCCCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCAAGATCCACGTTATTCGCCCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
Qy 2521 CCGAATCCCGAGATGTCGCTTTCTGGTATGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAATCCCGAGATGTCGCTTTCTGGTATGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGACGCTGCTGATGAAGCGCTGGAAGCGCGGAGCTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACGCTGCTGATGAAGCGCTGGAAGCGCGGAGCTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2700
Db 2641 AACAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2700
Qy 2701 GGATCC 2706
Db 2701 GGATCC 2706

RESULT 13
AF097412
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .6724
/organism="Expression vector pMBP-parallel1"
/mol_type="other DNA"
/db_xref="taxon:84479"
204. .1163
/gene="lacIq"
204. .1163
/gene="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAC83812.1"
/db_xref="GI:3983121"
/translation="MAELNYIPNRVAQAGKOSLLIGVATSSLAHPQSQIVAAIKS
RADQLGASVVSVMSRVSVEACAAVHNLQAVSGLIINYPDDQDAFAVEAACTNV
PALFLVSQDTPIINSIFSHEDGTRGLVHLVALGHQIALLAGPLSSVSARLRAGW
HKYLTNRQIPAEERDMSGSGFOOTMQLNEGIVPTAMLVANDOMALGAMRAITE
SLRVGADI SVVYDDTDSYCIPLPTIKODRLLGQTSVANDOMALGAMRAITE
LIPVSLVKETTLAENTQTASPRALADSLMARQVSRUESGQ"
1453. .1481
/note="tac; inducible promoter"
1528. .2994
/gene="male"
1528. .2994
/gene="male"
/codon_start=1
/transl_table=11
/product="maltose binding protein"
/protein_id="AAC83813.1"
/db_xref="GI:3983122"
/translation="MKIEGKLVIVINGDKYNGLAIEVGKKFEKDTGKIVTVEHPDKL
ESKFPQVAATGDPDIIFWHDRCFGYAGSGLLAIEITDKAFQDKLYFTWDVARYNG
KLIVPIAVEALSLINKDLLPNPPTWBIIPALDKELKAKGKSLMNFLOEYPTWP
LIAAGGTAFYENGKIDYKDVVDNAGAKAGLTFLVDLIKNKHNDTDSIAEAAE
NKGEMTINGPWSNIDTSKVNYGVTVLPFKGQSPFVGLSAGTNAASPNKEL
AKFLENLYLTDGELVAVKDKPLGAVALKSYEBELAKDPRIATMENAOKGEIMPN
PQMSAFVARTVINAASGRQTVDEALKDAQTSNNSSNNNNNNNNNNIGDITENLYF
QGAMPDFKGLRRRAQLVRPLSNLEPAVSRHVPSSLALAVLVLRQRDWENPGVTQLNRL
AAHPFPASWRNSEARTDRPSQQLSLNGEWQLGCGFGG"
2633. .2688
/gene="male"
/note="protein spacer region"
2689. .2709
/gene="male"
/note="Tobacco etch virus protease cleavage site"
2711. .2810
/gene="male"
misc_feature
misc_feature
misc_feature

gene /note="multiple cloning site"
3426..4356
/gene="bla"
promoter /product="beta-lactamase"
3426..3454
/gene="bla"
CDS 3496..4356
/gene="bla"
/codon_start=1
/transl_table=11
/protein_id="AAC83814.1"
/db_xref="GI:3983123"
/translation="MSIHFVRVALIPFAAFLPFAHPETLVKVKDAEDQLGARVGY
VELDNLASKILSFPRERSPMWSFQKLLCGAVLSRDVQGEQGRRIHYSDNDLVE
YSPVTEKHLTDGMTRELSLSAAITWSDNTANLLLTIGGPKELTAFLHNMGDHTRL
DRWPELNEAIPNDKRDTPVMAATLLKLTGELLTLASRQQLIDWMEADKVGPL
LRSLPAGWFIADKSGAGRGSGRIIAALPGDKPSRIIVIVITGSGQATMDERNRQIA
EIGASLIKHW"

ORIGIN

Query Match 81.1%; Score 2676.2; DB 12; Length 6724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAACCTTTCGGGGTATGCGCATGATAGCCCGGAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAACCTTTCGGGGTATGCGCATGATAGCCCGGAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACTGATATACGATGTCGAGAGTATGCCG 120
DB 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACTGATATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGCTTTCTGCGAAAA 180
DB 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGCTTTCTGCGAAAA 180
QY 181 CCGGGGAAAAAGTGGNAGCGGCGATGGCGGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
DB 181 CCGGGGAAAAAGTGGNAGCGGCGATGGCGGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
QY 241 AACAACTGCGGGGCAACAGTGTGCTGATTTGGGGTTCGCCACCTCCAGCTTGGCCCTCG 300
DB 241 AACAACTGCGGGGCAACAGTGTGCTGATTTGGGGTTCGCCACCTCCAGCTTGGCCCTCG 300
QY 301 ACGCGCGTTCGCAAAATTTGTCGGGCGATTAATCTCGCGCGGATCAACTGGGTGGCCAGCG 360
DB 301 ACGCGCGTTCGCAAAATTTGTCGGGCGATTAATCTCGCGCGGATCAACTGGGTGGCCAGCG 360
QY 361 TGGTGTGTGATGTAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCACAATC 420
DB 361 TGGTGTGTGATGTAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGGATGCCA 480
DB 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGTGCCTGCATTAATTTCCGCGGTTATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGGAAGTGCCTGCATTAATTTCCGCGGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGATCGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGATCGGCGGTGGAGCATC 600
QY 601 TGGTGCATTTGGTTCACCGCAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTGCATTTGGTTCACCGCAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CCGGTCTCGGTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CCGGTCTCGGTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAAACGGGAAGCGGACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGTA 780

DB 721 CCGAAACGGGAAGCGGACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCCAACGATCGCGTGGCGCGCAA 840
DB 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCCAACGATCGCGTGGCGCGCAA 840
QY 841 TCGCGCCCATTAACGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 841 TCGCGCCCATTAACGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGGAAGACAGCTCATGTTATATATCCCGCGTGAACACCATCAAAACGAGATTTC 960
DB 901 ACGATACCGGAAGACAGCTCATGTTATATATCCCGCGTGAACACCATCAAAACGAGATTTC 960
QY 961 GCCTGCTCGGGGCAAAACACGAGCTGGACCGCTTGTGTCACCTCTCTCAGGCGCAGCGGTGA 1020
DB 961 GCCTGCTCGGGGCAAAACACGAGCTGGACCGCTTGTGTCACCTCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGGCAATCAGTGTGTCCTCTCCTCAGTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
DB 1021 AGGGCAATCAGTGTGTCCTCTCCTCAGTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCACGACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCACGACAGGTTT 1140
QY 1141 CCGGACTCGAAAGCGGCGAGTGGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
DB 1141 CCGGACTCGAAAGCGGCGAGTGGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGCAGCTTATCATCGATGTCACGCTGCACCAATGTTCTGGCG 1260
DB 1201 GCACAATTTCTCATGTTTGCAGCTTATCATCGATGTCACGCTGCACCAATGTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGTGCAGTGCCTGTAATCACTGATATTCG 1320
DB 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGTGCAGTGCCTGTAATCACTGATATTCG 1320
QY 1321 TGTCTCTCAAGCGGCACTCCCGCTTCGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
DB 1321 TGTCTCTCAAGCGGCACTCCCGCTTCGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATAATTTCTGAAATGAGCTGTGACAAATTAATCATCGCTCGTATAATGTGTGA 1440
DB 1381 CTGGCAATAATTTCTGAAATGAGCTGTGACAAATTAATCATCGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTCACACAGGAACACCGCAGTCCGTTTAGTGTTCACGA 1500
DB 1441 ATTGTGAGCGGATAACAATTTTCACACAGGAACACCGCAGTCCGTTTAGTGTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGNTAACTGGAAGAGAAAATTCCTCCACAGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGNTAACTGGAAGAGAAAATTCCTCCACAGTT 1680
QY 1681 GGGGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTCGTGGCTAC 1740
DB 1681 GGGGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTCGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGCAAGCGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGCAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTAGCTTACAAACGGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860

Db 1801 CCGTTTACCTGGGATGCGCTACGTTACAAACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATATAACAAGATCTGCTGCGACCGCCCAAAAACCTGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATATAACAAGATCTGCTGCGAACCGCCCAAAAACCTGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGGCAAGTACGACATTAAAGAGCGTGGCGTGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTAAAGAGCGTGGCGTGATAAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAAACAAACACATGATGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAAACAAACACATGATGACACCGGATTAC 2160
QY 2161 TCCATCGCAGAAGTGCCTTTAATAAGGCGAAACAGCGATGACCATCAACCGCGCGTGC 2220
Db 2161 TCCATCGCAGAAGTGCCTTTAATAAGGCGAAACAGCGATGACCATCAACCGCGCGTGC 2220
QY 2221 GCATGTCGAACATCGACACGACCAAGTAAATGATGTAACGGTCTGCGGACCTTC 2280
Db 2221 GCATGTCGAACATCGACACGACCAAGTAAATGATGTAACGGTCTGCGGACCTTC 2280
QY 2281 AAGGTCGAACATCGACACGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2340
Db 2281 AAGGTCGAACATCGACACGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2340
QY 2341 CCGAAACAAAGAGTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATCAAGGTCG 2400
Db 2341 CCGAAACAAAGAGTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATCAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAAGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAAGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGGTGCGTGGTGGTGGTGGTGGTGGTGGTGG 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGGTGCGTGGTGGTGGTGGTGGTGGTGGTGG 2580
QY 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGGAAGACGGCGAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGGAAGACGGCGAGACTTAATTCGAGCTCG 2640
QY 2641 AACACACAAACAATAACAATAACAACACCTCGGGATCGA 2681
Db 2641 AACACACAAACAATAACAATAACAACCTCGGGATCGA 2681

RESULT 14
LOCUS AX378208
DEFINITION Sequence 194 from Patent W00206306.
ACCESSION AX378208
VERSION AX378208.1 GI:19574061
KEYWORDS .
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1
AUTHORS Yan, R., Tomasselli, A.G., Gurney, M.E., Emmons, T.L., Bienkowski, M.J. and Heinrikson, R.L.
TITLE Substrates and assays for _g(b)-secretase activity

JOURNAL Patent: WO 0206306-A 194 24-JAN-2002;
FEATURES PHARMACIA & UPJOHN COMPANY (US)
source Location/Qualifiers
1. .6806
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="fusion protein comprising a maltose binding protein with 125 amino"

ORIGIN

Query Match
Best Local Similarity 79.9%; Score 2636.8; DB 6; Length 6806;
Matches 2638; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGCGTATGGCATGATAGCCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGCGCAAAACCTTTCCGCGTATGGCATGATAGCCGCCGGAAGAGA 60
QY 61 GTCATTTCCAGGTTGTAATGTGAACACAGTAACGTTTATACGATGTCGACAGATATGCCG 120
Db 61 GTCATTTCCAGGTTGTAATGTGAACACAGTAACGTTTATACGATGTCGACAGATATGCCG 120
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACAGGCGCAGCCAGCTTTCTGCGAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACAGGCGCAGCCAGCTTTCTGCGAAAA 180
QY 181 CCGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAACATTTCCCAACCGCGTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCGCAACAGTCGTTGCTGATTTGGCGTTGGCACCTCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCGCAACAGTCGTTGCTGATTTGGCGTTGGCACCTCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCGAAATTTGTCCGCGCGATTAATTAATCTCGCGCGGATCAACTGGGTGCCAGG 360
Db 301 ACGCGCGTCCGAAATTTGTCCGCGCGATTAATTAATCTCGCGCGGATCAACTGGGTGCCAGG 360
QY 361 TGTGTTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGCGGTGCAAAATC 420
Db 361 TGTGTTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGCGGTGCAAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATTAATCTCGCGCGGATCAACTGGGTGCCAGG 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATTAATCTCGCGCGGATCAACTGGGTGCCAGG 480
QY 481 TTGCTGTGAGAGCTGCTGCACTAATGTTCCGCGGTTATTTCTTGATGTTCTTGACCCAGA 540
Db 481 TTGCTGTGAGAGCTGCTGCACTAATGTTCCGCGGTTATTTCTTGATGTTCTTGACCCAGA 540
QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGCGTGGAGCATC 600
QY 601 TGTGTCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCGCG 660
Db 601 TGTGTCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCGCG 660
QY 661 CCGCTCTGCTGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGCTCTGCTGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAAACGGGAAGCGGACCTGGAGTGCATGTCGCGTTTTCACAAACCAACATGAAATGCTGA 780
Db 721 CCGAAACGGGAAGCGGACCTGGAGTGCATGTCGCGTTTTCACAAACCAACATGAAATGCTGA 780
QY 781 ATGAGGCGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATGATGCGCTGGCGCAA 840
Db 781 ATGAGGCGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATGATGCGCTGGCGCAA 840
QY 841 TCGCGCCCATTAACCGAGTCCGCGCTGGTGGCGGATATCTCGTACTGCGGATACG 900

Db 841 TCGCGCCATTACCGAGTCCGGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAAAACCGAGCTGGACCGTTCCTGCTCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACCGAGCTGGACCGTTCCTGCTCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTCGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTCGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Qy 1081 GCGAAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGCGACGACAGGTTT 1140
Db 1081 GCGAAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGCGACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGACGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGACGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCGCGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCGCGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAATCACTGCTAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAATCACTGCTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCGTAATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCGTAATGTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTAGGTGTTTTTCAGCA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTAGGTGTTTTTCAGCA 1500
Qy 1501 GCATTTCACCAAGGACCATAGATTATGAACCTGAAGAAAGGTAACCTGGTAACTCTGG 1560
Db 1501 GCATTTCACCAAGGACCATAGATTATGAACCTGAAGAAAGGTAACCTGGTAACTCTGG 1560
Qy 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 GCGGCAATCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTGGTGGCTAC 1740
Db 1681 GCGGCAATCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCCGCAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCCGCAAGCGTTCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCCGTAGCTTACACGCGCAAGCTGATTGCTTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCCGTAGCTTACACGCGCAAGCTGATTGCTTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTTTATCAAGATCTGCTGCGGAAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATCAAGATCTGCTGCGGAAACCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980

Qy 1981 CTGCAAGAACCGTATTCTACCTGGCCGCTGATTCTGCTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTATTCTACCTGGCCGCTGATTCTGCTGACGGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAAACGCGAAGTAGCATTAAAGACCTGGGGCTGGATTAACGCTGGGCGGAAAGCG 2100
Db 2041 TATGAAAACGCGAAGTAGCATTAAAGACCTGGGGCTGGATTAACGCTGGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Qy 2221 GCATGGTCCAACTCGACACGACCAAGTGAATTAATGGTGAACCGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACGACCAAGTGAATTAATGGTGAACCGTACTGCCGACCTTC 2280
Qy 2281 AAGGTCACACCATCAAAACCGTTCTGGCGTGTCTGAGCGAGGTATTAACGCGCCCAAGT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTCTGGCGTGTCTGAGCGAGGTATTAACGCGCCCAAGT 2340
Qy 2341 CCGAACAAGAGCTGGCGAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCGAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Qy 2401 GAAGGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATTCGCCCAACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTCGCCCAACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGCTCGCTTTCTGGTATGCGCGTGTCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGCTCGCTTTCTGGTATGCGCGTGTCTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGCTGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640

RESULT 15
AF031088 6706 bp DNA circular SYN 07-MAY-2001
LOCUS Shuttle vector pMAL-pIII, complete sequence.
DEFINITION AF031088
ACCESSION AF031088
VERSION AF031088.1 GI:2623821
KEYWORDS Shuttle vector pMAL-pIII
SOURCE Shuttle vector pMAL-pIII
ORGANISM Shuttle vector pMAL-pIII
Other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6706)
AUTHORS Zwick,M.B., Bonnycastle,L.L., Noren,K.A., Venturini,S., Leong,E.,
Barbas,C.F. III, Noren,C.J. and Scott,J.K.
TITLE The maltose-binding protein as a scaffold for monovalent display of
peptides derived from phage libraries
JOURNAL Anal. Biochem. 264 (1), 87-97 (1998)
MEDLINE 99002881
PUBMED 9784192
REFERENCE 2 (bases 1 to 6706)
AUTHORS Noren,K.A. and Noren,C.J.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road,
Beverly, MA 01915, USA
FEATURES location/Qualifiers
source 1..6706
/organism="Shuttle vector pMAL-pIII"
/mol_type="genomic DNA"

```

/db xref="taxon:69327"
/notes="derived from commercially available vector pMal-p2
(New England Biolabs)"
15..1163
/gene="lacI"
/promoter
15..46
/gene="lacI"
81..1163
/gene="lacI"
/notes="transcriptional repressor"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAB86558.1"
/db_xref="GI:2623822"
/translation="MKPVTLYDVAEYAGVSQYTVSRVNVQASHVSAKTRKVEAAMAE
LNYINRVAQAGSLGIVATSSLAHAPSQIVAAIKSRADQLGASVVSVMVERS
GVKCAKAAVHNLIAQVRSGLIINYPLDDQDAIAVEAACTNVPALFLVSDQTPINSII
FSHEDGRLGVHVALGHQIALLAGPLSSVARSRLAGWHKYLTRNQIOPIABREG
DMSAGSFQOTQMINEGIVPTAMLVANDQMALGAMRAITSGSLRVGADISVVGVDYT
EDSSCVIPPLTTIKODPRLGOTSVDRLQLSQQAVKGNQLLPSLVKRTTLAPNT
QTASPRALADSLMQLARQVSRLESQ"
1406..1433
/promoter
/notes="Plac promoter"
1528..2976
/gene="male-lacZa fusion"
/notes="E. coli gene fusion"
1528..2976
/gene="male-lacZa fusion"
/notes="MBP-lacZa; MBP modified to include phage M13
secretion signal sequence"
/codon_start=1
/transl_table=11
/product="maltose binding protein-lacZ alpha peptide
fusion protein precursor"
/protein_id="AAB86559.1"
/db_xref="GI:2623823"
/translation="MKKLLFAIPLVVPFISHSADMKBEGKLVIWINGDKYNGLAEV
GKFFKDTGIVKTVVHPDKLEKFPQVATGDDPDIIFWAHDFRGVYQSGLLAEITP
DKAFQDLYPTWADRVNKGKLIAYIAVEALSIIYNKDLIPNPKTWEEIIPALDKEL
KAKGSAALMFLNLOBEYFWPLIADGGVAFKVGKDYIDKVDGVDNAGAKAGLFLVD
LIFKNHMADYDIAEAAFNKGETAMTINGPWAWSNIDTSKVNGVTVLPTFKGQPS
KPFVGLSAGINASPNKELAKFLEFLNLLTDEGLEAVNKDKPLIGVALKSYEBELAK
DPRIATWENAKGIMPNIPQMSAFWYAVTAVINASGRQTVDEALKDAQTNSSN
NNNNNNNLGIEISFSGSRVDLQASLALAVVLQRDWNENPGVTQLNRLAAHPFP
ASWRNSEARTDRPSQQLRSLNGEWQLGCFG"
1528..1581
/gene="male-lacZa fusion"
/notes="M13 secretion signal sequence"
1582..2973
/gene="male-lacZa fusion"
/product="maltose binding protein-lacZ alpha peptide
fusion protein"
3478..4338
/gene="bla"
/notes="R-plasmid ampicillin resistance gene"
3478..4338
/gene="bla"
/notes="penicillinase"
/codon_start=1
/transl_table=11
/product="RTEM-1 beta-lactamase"
/protein_id="AAB86560.1"
/db_xref="GI:2623824"
/translation="MSIQHFRVALIPFFAAFLPFAHPETLVKVKDAEDQLGARVGY
IELDNSKKILESFRPERSPMSTFKVLLCGAVLSRVDAQIQQLGRRRIHYSNDLVE
YSPVTEKHLTDGMTRVRLTSCGAAITMSDNTAANLLLTIGGPKELTAFLNHMGPHVRL
DRWPELNEATPNDERDTTPVAMATTLKLLTGLLTLASRQLDLWMEADKVAQPL
LRSLALPAGWFTADSKSGRSGRIIAALGPDGKPSRIVVIYITGDSQATMDERNQIA
EIGASLIKHW"
4381..4893
/standard_name="M13 origin (+ -)"

```

ORIGIN

```

Query Match
Best Local Similarity 79.7%; Score 2631; DB 12; Length 6706;
Matches 2704; Conservative 0; Mismatches 5; Indels 60; Gaps 1;

QY 1 CGGACACCATCGAATGGTGTGCAAAACCTTTTCGCGTATGGCATGATAGCCCGGAAGAGA 60
DB 1 CGGACACCATCGAATGGTGTGCAAAACCTTTTCGCGTATGGCATGATAGCCCGGAAGAGA 60

QY 61 GTCAAATTCAGGCTGTGAATGTGAACCAAGTAACTTATACGATGTGCAGAGTATGCCG 120
DB 61 GTCAAATTCAGGCTGTGAATGTGAACCAAGTAACTTATACGATGTGCAGAGTATGCCG 120

QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGTAACACGAGCCAGCCAGTCTTCTCGAAAA 180
DB 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGTAACACGAGCCAGCCAGTCTTCTCGAAAA 180

QY 181 CGCGGGAAGAAGTGAAGCGCGCATGGCGAGCTGAATTAATTCCTCCACCGCGTGGCAC 240
DB 181 CGCGGGAAGAAGTGAAGCGCGCATGGCGAGCTGAATTAATTCCTCCACCGCGTGGCAC 240

QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300

QY 301 ACGCGCGTGTGCAAAATTTGTGCGCGCATTAATACTCGCGCGCATCAACTGGGTGCCACGC 360
DB 301 ACGCGCGTGTGCAAAATTTGTGCGCGCATTAATACTCGCGCGCATCAACTGGGTGCCACGC 360

QY 361 TGGTGGTGTGATGGTAGAAGCGCGCTGCAAGCGCTGAAGCGCTGAAGCGCGGTGCACATC 420
DB 361 TGGTGGTGTGATGGTAGAAGCGCGCTGCAAGCGCTGAAGCGCGGTGCACATC 420

QY 421 TTCTCGCGCAACGCGTCAGTGGCTGATCAATTAATCACTCCGCTGGATGACAGAGTGCCA 480
DB 421 TTCTCGCGCAACGCGTCAGTGGCTGATCAATTAATCACTCCGCTGGATGACAGAGTGCCA 480

QY 481 TTGCTGTGGAGCTGCTGCACCTAATGTTCCGCGCTTATTTCTTGATGTCTCTGACAGA 540
DB 481 TTGCTGTGGAGCTGCTGCACCTAATGTTCCGCGCTTATTTCTTGATGTCTCTGACAGA 540

QY 541 CACCCATCAACAGTATTTCTCCCATGAACAGCGTACGCGACTGGCGTGGGAGCATC 600
DB 541 CACCCATCAACAGTATTTCTCCCATGAACAGCGTACGCGACTGGCGTGGGAGCATC 600

QY 601 TGTGCGATTTGGGTACACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660
DB 601 TGTGCGATTTGGGTACACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660

QY 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAACGGGAAGCGCATGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTCA 780
DB 721 CGGAACGGGAAGCGCATGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTCA 780

QY 781 ATGAGGGCATCTGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGGCGTGGGCGCAA 840
DB 781 ATGAGGGCATCTGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGGCGTGGGCGCAA 840

QY 841 TGGCGGCCATTACCGAGTCCGGCTGCGGCTGGTGGCGATATCTCGGTAGTGGGATAGC 900
DB 841 TGGCGGCCATTACCGAGTCCGGCTGCGGCTGGTGGCGATATCTCGGTAGTGGGATAGC 900

QY 901 ACGATACCAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
DB 901 ACGATACCAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 960

QY 961 GCTGCTGGGGCAAAACAGCGTGGAGCGCTTCTGCTCACTCTCAGGGCCAGGGGTGA 1020
DB 961 GCTGCTGGGGCAAAACAGCGTGGAGCGCTTCTGCTCACTCTCAGGGCCAGGGGTGA 1020

```

Qy	1021	AGGCAATCAGCTGTTGCCGTCCTCACTGTTGAAAGAAAAACAACCTTGGCCGCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCGTCCTCACTGTTGAAAGAAAAACAACCTTGGCCGCCAATA	1080
	1081	CGAAAAACGCTCTCCCGCGGCTTGGCCGATTCATTAATGCAAGTGGCACGACAGGTTT	1140
Db	1081	CGAAAAACGCTCTCCCGCGGCTTGGCCGATTCATTAATGCAAGTGGCACGACAGGTTT	1140
	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Qy	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG	1260
	1261	TCAGCAGCCCATCGAAAGCTGTGTATGTGGTGTGCAGGTGCTAAATCACTGCATAATTCG	1320
Qy	1261	TCAGCAGCCCATCGAAAGCTGTGTATGTGGTGTGCAGGTGCTAAATCACTGCATAATTCG	1320
	1321	TGTCGCTCAAGCGCACTCCCGTTCTCGATAATGTTTTTTCGCGCAGCATATAACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACTCCCGTTCTCGATAATGTTTTTTCGCGCAGCATATAACGGTT	1380
	1381	CTGCAAAATTTCTGAAATGAGCTGTTGCACAAATTAATCATCTGGCTCGTATAATGTGTGA	1440
Qy	1381	CTGCAAAATTTCTGAAATGAGCTGTTGCACAAATTAATCATCTGGCTCGTATAATGTGTGA	1440
	1441	ATTGTAGCGGATAAACAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTACGA	1500
Db	1441	ATTGTAGCGGATAAACAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTACGA	1500
	1501	GCATTTCAACCAAGGACCATAG-----	1524
Qy	1501	GCATTTCAACCAAGGACCATAGCATATAAAAAATTAATTCGCAATTCCTTTAGTG	1560
	1525	-----ATTATGAAACTGAAGAGTAACTGTAATCTGG	1560
Db	1561	GTAACCGCGATAAAGCTATAACCGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT	1620
	1621	ATTAAACCGCGATAAAGCTATAACCGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT	1680
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATATACTGGAAGAGAAATTCACACAGTT	1680
	1681	ACCGGAATTAAGTCAACCGTTGAGCATCCGATATACTGGAAGAGAAATTCACACAGTT	1740
Qy	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGGTGCTAC	1740
	1741	CGCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGGTGCTAC	1800
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGCAACAGCTGTAT	1800
	1801	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGCAACAGCTGTAT	1860
Qy	1801	CCGTTTACCTGGGATGGCTGACGTTAACAACGCAAGCTGATGCTTACCCGATCGCTGTT	1860
	1861	CCGTTTACCTGGGATGGCTGACGTTAACAACGCAAGCTGATGCTTACCCGATCGCTGTT	1920
Qy	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGACACCGCCGCAAAAACCTGGAA	1920
	1921	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGACACCGCCGCAAAAACCTGGAA	1980
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTTCAAC	1980
	1981	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTTCAAC	2040
Qy	1981	CTGCAAGAACCGTATCTTCACTGCGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG	2040
	2041	CTGCAAGAACCGTATCTTCACTGCGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG	2100

Qy	2041	TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGATTAACGCTGGCGGAAGCG	2100
Db	2101	TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGATTAACGCTGGCGGAAGCG	2160
Qy	2101	GGTCTGACCTTCCTGGTTGACCTGATTAAAAAACAACACATGAATGAGACACCGGATTAC	2160
Db	2161	GGTCTGACCTTCCTGGTTGACCTGATTAAAAAACAACACATGAATGAGACACCGGATTAC	2220
Qy	2161	TCCATCGCAGAGCTGCTTTTAATAAGGGCGAAACAGCGATGACCATCAACGGCGCCGTGG	2220
Db	2221	TCCATCGCAGAGCTGCTTTTAATAAGGGCGAAACAGCGATGACCATCAACGGCGCCGTGG	2280
Qy	2221	GCATGTTCCAACATCGCACACAGCAAAAGTCAAATTAATGGTGTAAACGGTACTGCCGACCTTC	2280
Db	2281	GCATGTTCCAACATCGCACACAGCAAAAGTCAAATTAATGGTGTAAACGGTACTGCCGACCTTC	2340
Qy	2281	AAGGGTCAACCATCAAAACCGTTGTTGGTGCTGAGCGCAGGTATTAAACGGCGCCAGT	2340
Db	2341	AAGGGTCAACCATCAAAACCGTTGTTGGTGCTGAGCGCAGGTATTAAACGGCGCCAGT	2400
Qy	2341	CCGAACAAAGAGCTGGCAAAAGATTCCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
Db	2401	CCGAACAAAGAGCTGGCAAAAGATTCCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2460
Qy	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Db	2461	GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAG	2520
Qy	2461	TTGGCGAAGAGTCCACGTATTGCGGCCACCATGGAAACCGCCAGAAAGTGAAATCATG	2520
Db	2521	TTGGCGAAGAGTCCACGTATTGCGGCCACCATGGAAACCGCCAGAAAGTGAAATCATG	2580
Qy	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCCGTGCGTACTCGGGTGATCAACGCC	2580
Db	2581	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCCGTGCGTACTCGGGTGATCAACGCC	2640
Qy	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCG	2640
Db	2641	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCG	2700
Qy	2641	AACAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2701	ARCAACAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2760
Qy	2701	GGATCTCTT 2709	
Db	2761	GGATCTCTT 2769	

Search completed: August 1, 2005, 17:21:40
Job time : 9286 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 14:24:04 ; Search time 1094.6 Seconds

(without alignments)
17846.836 Million cell updates/sec

Title: US-09-765-555B-17

Perfect score: 3300

Sequence: 1 ccgacacatcgatgtgc.....acgacgttcgcgactaagct 3300

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	4 AAD11591	Aad11591 Partial s
2	3232.8	98.0	3300	4 AAD11592	Aad11592 Partial s
3	3215.2	97.4	3300	4 AAD11588	Aad11588 Partial s
4	3215.2	97.4	3300	4 AAD11589	Aad11589 Partial s
5	3207.2	97.2	3300	4 AAD11590	Aad11590 Partial s
6	2705.8	82.0	8101	4 AAD25681	Aad25681 Nucleotid
7	2705.8	82.0	8101	6 ABL53238	ABL53238 Nucleotid
8	2690.4	81.5	7259	12 ADO23608	ADO23608 DNA encod
9	2690.4	81.5	7322	12 ADO23613	ADO23613 DNA encod
10	2690.4	81.5	7352	12 ADO23603	ADO23603 DNA encod
11	2690.4	81.5	7370	12 ADO23598	ADO23598 DNA encod
12	2690.4	81.5	7370	12 ADO23639	ADO23639 DNA encod
13	2690.4	81.5	7370	12 ADO23649	ADO23649 DNA encod
14	2690.4	81.5	7370	12 ADO23644	ADO23644 DNA encod
15	2690.4	81.5	7403	12 ADO23594	ADO23594 DNA encod
16	2690.4	81.5	7442	12 ADO23590	ADO23590 DNA encod
17	2690.4	81.5	7478	12 ADO23584	ADO23584 DNA encod
18	2688.8	81.5	7112	12 ADO23618	ADO23618 DNA encod
19	2636.8	79.9	6806	6 ABL49925	ABL49925 Maltose b
20	2605.4	79.0	7553	12 ADO23588	ADO23588 DNA encod

21	1609.4	48.8	5558	12	ADL90419	Adl90419 Clostridi
22	1474	44.7	4700	9	ACF06053	Acf06053 Vector pJ
23	1472.4	44.6	4700	12	ADL18670	Adl18670 Vector pJ
24	1462.4	44.3	4920	10	ADH73599	Adh73599 Plasmid p
25	1462.4	44.3	4935	10	ADH73598	Adh73598 Plasmid p
26	1462.4	44.3	4945	10	ADH73600	Adh73600 Plasmid p
27	1462.4	44.3	4951	10	ADH73601	Adh73601 Plasmid p
28	1418.2	43.0	5903	8	ABZ23939	Abz23939 Nucleotid
29	1417.6	43.0	1511	6	ABQ73200	Abq73200 E. coli t
30	1417.6	43.0	5926	2	AAV32977	Aav32977 Tn7 donor
31	1417.6	43.0	5926	6	AAD45059	Aad45059 Transposo
32	1417.6	43.0	5926	12	ADG46817	Adg46817 Donor pla
33	1362	41.3	5201	12	ADL72229	Adl72229 DNA seque
34	1362	41.3	5201	12	ADL72228	Adl72228 DNA seque
35	1238.4	37.5	5024	9	ACF06051	Acf06051 Plasmid p
36	1238.4	37.5	5024	12	ADL18668	Adl18668 Plasmid p
37	1237.2	37.5	1922	6	ABQ73089	Abq73089 MBP and z
38	1237.2	37.5	1922	10	ADF83622	Adf83622 MBP-human
39	1204	36.5	4557	2	AAT90491	Aat90491 Vector pl
40	1204	36.5	4969	12	ADO33522	Ado33522 Expressio
41	1204	36.5	5006	10	ADE24112	Ade24112 plasmid p
42	1204	36.5	5064	12	ADO33532	Ado33532 Vector pG
43	1204	36.5	5082	12	ADO33531	Ado33531 Vector pG
44	1204	36.5	6259	6	AAO29720	Aao29720 Plasmid p
45	1204	36.5	6823	3	AAC55459	Aac55459 Destinati

ALIGNMENTS

RESULT 1		AAD11591		AAD11591 standard; DNA; 3300 BP.	
ID	AAD11591	AC	AAD11591;	AC	AAD11591;
XX		AC	AAD11591;	AC	AAD11591;
XX		AC	AAD11591;	AC	AAD11591;
DT	24-SEP-2001	(first entry)			
XX		Partial sequence of pMal-m4 and ZFPm4 DNA.			
DE		Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;			
KW		modulation; plant technology; agriculture; ds.			
XX		Unidentified.			
OS		Unidentified.			
XX	Key	Location/Qualifiers			
FT	CDS	2719..3270			
FT		/*tag= a			
FT		/product= "ZFPm4 protein"			
FT		/transl_except= (pos:3046..3048, aa:Ser)			
FT		/note= "CDS does not include start and stop codon"			
FT		/partial			
FT	primer_bind	2740..2790			
FT		/*tag= b			
FT		/bound_moiety= "F1-f2 primer"			
FT	primer_bind	2770..2850			
FT		/*tag= c			
FT		/bound_moiety= "F1-f1 primer"			
FT	primer_bind	2824..2889			
FT		/*tag= d			
FT		/bound_moiety= "F2-b primer"			
FT	primer_bind	2867..2940			
FT		/*tag= e			
FT		/bound_moiety= "F2-f primer"			
FT	primer_bind	2916..2973			
FT		/*tag= f			
FT		/bound_moiety= "F3-b1 primer"			
FT	primer_bind	2953..3021			
FT		/*tag= g			
FT		/bound_moiety= "F3-b2 primer"			
FT	primer_bind	2992..3042			
FT		/*tag= h			
FT		/bound_moiety= "F4-f2 primer"			

```
FT primer_bind 3022..3102
FT /tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /tag= m
FT /bound_moiety= "F6-b2 primer"
XX WO200152620-A2.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 19-JAN-2001; 2001WO-US001817.
PF
XX
XX 21-JAN-2000; 2000US-0177468P.
PR
XX
XX 21-JUL-2000; 2000US-00620897.
PR
XX
XX (SCRI ) SCRIPPS RES INST.
PA
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CF, Stege JT, Guan X, Dalmia B;
PI
XX WPI: 2001-465325/50.
XX
XX P-PSDB; AAE06004.
DR
XX
XX New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
XX Example 4; Page 145-147; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m4 and ZFPm4 DNA
XX
SQ Sequence 3300 BP; 849 A; 865 C; 880 G; 706 T; 0 U; 0 Other;

Query Match 100.0%; Score 3300; DB 4; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGAACACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60
DB 1 CCGACACCATCGAATGGTGAACACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60
QY 61 GTCAATTGAGGTGGTGAATGTGAACACGTAACGTTATACGATGTGCGAGAGTATGCCG 120
DB 61 GTCAATTGAGGTGGTGAATGTGAACACGTAACGTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGCTTTCGCGAAAA 180
DB 121 GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGCTTTCGCGAAAA 180
QY 181 CGCGGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAAATTACATCCCAACCGCGTGGCAC 240
DB 181 CGCGGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAAATTACATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCGTCCCAATTTGTCGGCGGATTTAAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGGCGCGTCCCAATTTGTCGGCGGATTTAAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTAACGAAAGCGCGTCGAAGCTCTGTAAGCGCGCGGTGCACAATC 420
DB 361 TGGTGGTGTGATGTAACGAAAGCGCGTCGAAGCTCTGTAAGCGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATCTCCGCTGGATGACACGAGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATCTCCGCTGGATGACACGAGTGCCA 480
QY 481 TTGCTGTGGAGCTGCTGCACTAAATGTTCCGGGGTTATTTCTTGATGTTCTTGACACAGA 540
DB 481 TTGCTGTGGAGCTGCTGCACTAAATGTTCCGGGGTTATTTCTTGATGTTCTTGACACAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGATTTGGGTACCAGCAAAATCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
DB 601 TGGTCGATTTGGGTACCAGCAAAATCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGGTCCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGGTCCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTCAACAAACCATCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTCAACAAACCATCAAAATGCTGA 780
QY 781 ATGAGGCGATCGTTCCTCCACTGCGATGTTGTTGCCAACGATGAGTGGCGCTGGGCGCAA 840
DB 781 ATGAGGCGATCGTTCCTCCACTGCGATGTTGTTGCCAACGATGAGTGGCGCTGGGCGCAA 840
QY 841 TGGCGGCCATTACCGAGTCCGGCTCGCGTTCGGTGGGATATCTCGTAGTGGGATACG 900
DB 841 TGGCGGCCATTACCGAGTCCGGCTCGCGTTCGGTGGGATATCTCGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGTCAATGTTATATCCGCGGTTAACCCACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGTCAATGTTATATCCGCGGTTAACCCACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAACCGAGTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGA 1020
DB 961 GCCTGCTGGGGCAACCGAGTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTTCAATTAATGAGCTGGGCAACGAGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTTCAATTAATGAGCTGGGCAACGAGTTT 1140
QY 1141 CCGGCTGGAAGAGCGGCGAGTGGCGCAACGCAATTAATGAGCTGGGCAACGAGTTT 1200
DB 1141 CCGGCTGGAAGAGCGGCGAGTGGCGCAACGCAATTAATGAGCTGGGCAACGAGTTT 1200
QY 1201 GCACAAATCTCTGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGGG 1260
DB 1201 GCACAAATCTCTGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGGG 1260
QY 1261 TCAGGCGAGCCATCGGAAGCTGTTGGCTGCGAGTGGTAAATCACTGCAATATTCG 1320
DB 1261 TCAGGCGAGCCATCGGAAGCTGTTGGCTGCGAGTGGTAAATCACTGCAATATTCG 1320
```

Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCG 1320
Qy 1321 TGTGCGTCAAGGGGCACTCCCGTTCTGGATAAATGTTTTTGGCGCGACATCATAAACGGTT 1380
Db 1321 TGTGCGTCAAGGGGCACTCCCGTTCTGGATAAATGTTTTTGGCGCGACATCATAAACGGTT 1380
Qy 1381 CTGGCAAAATATCTGAAATAGCTGTGACAATTAATCATCGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATAGCTGTGACAATTAATCATCGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTTCAACAGAGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTCAACAGAGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Qy 1501 GCACCTTACCACCAACAGGACCATAGATTATGAAAACTGAAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTACCACCAACAGGACCATAGATTATGAAAACTGAAAGAGGTAAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAAATTAAGTCAACCGTTGAGATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAAATTAAGTCAACCGTTGAGATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCAACGACGCTTTGGTGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCAACGACGCTTTGGTGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGTGGCTGAATCAACCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAATCAACCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGGCGCTAGCTTACAACGGCAAGCTGATTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGGCGCTAGCTTACAACGGCAAGCTGATTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAAAGATCTGCTCGGAAACCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAAAGATCTGCTCGGAAACCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGCTGACGGGGTTATCGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGCTGACGGGGTTATCGCTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAACACATGAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGCAAGAGCTGCCCTTTAATAAGCGGAAACAGCGCATGACCATCAACCGCCGCTGG 2220
Db 2161 TCCATCGCAAGAGCTGCCCTTTAATAAGCGGAAACAGCGCATGACCATCAACCGCCGCTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACGCAAGTGAATTTATGGTGAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACGCAAGTGAATTTATGGTGAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACCGCCGAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACCGCCGAGT 2340
Qy 2341 CCGAACAAAGAGCTGCGCAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGCGCAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCTG 2400

Qy 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGGAAAAGATCCACGATTATGGCCCAACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGAAAAGATCCACGATTATGGCCCAACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCCGTCTGTAATCTGAGAGCTC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCCGTCTGTAATCTGAGAGCTC 2580
Qy 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCCAGAAAGGTGAAATTC 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCCAGAAAGGTGAAATTC 2640
Qy 2641 AACAAACAACAATAAACAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAACAACAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCCTCTTCTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAAGCCCTATGCTGT 2760
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGTAGTCTTCAAGCAGAGCAGCTCCCTGGTGGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAGTCTTCAAGCAGAGCAGCTCCCTGGTGGCCACACAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGAGCAGCAG 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGAGCAGCAG 2880
Qy 2881 CTGTCGCCCATCAACGCACTCATATCTGGCGAGAAAGCCATACAAATGTCCAGAAATGTG 2940
Db 2881 CTGTCGCCCATCAACGCACTCATATCTGGCGAGAAAGCCATACAAATGTCCAGAAATGTG 2940
Qy 2941 AAGTCTTTCAGTGAATGTGATCTTGGAGGACCAACAGTACTCAACCCGGGAGAAG 3000
Db 2941 AAGTCTTTCAGTGAATGTGATCTTGGAGGACCAACAGTACTCAACCCGGGAGAAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCCTGGTGGCCAC 3060
Qy 3061 CAGCGTACCCACACCGGGTCAAAAAACCGTATAAATGCCAGAGTGGCGGAAAATCTTTTAG 3120
Db 3061 CAGCGTACCCACACCGGGTCAAAAAACCGTATAAATGCCAGAGTGGCGGAAAATCTTTTAG 3120
Qy 3121 CGCAGCGATAACCTGGTGGCCATCAACGCACTCATATCTGGCGAGAAGCCATACAAATGT 3180
Db 3121 CGCAGCGATAACCTGGTGGCCATCAACGCACTCATATCTGGCGAGAAGCCATACAAATGT 3180
Qy 3181 CCAGAAATGCGCAAGTCTTCTCAACTTTCAGGCCATTTGGTCCGTCACCAAGTACTCAC 3240
Db 3181 CCAGAAATGCGCAAGTCTTCTCAACTTTCAGGCCATTTGGTCCGTCACCAAGTACTCAC 3240
Qy 3241 ACCCGTAAAAAACTAGTGGCCAGCGCCAGTACCCGCTAGAGGTTCCGAGCTACGCT 3300
Db 3241 ACCCGTAAAAAACTAGTGGCCAGCGCCAGTACCCGCTAGAGGTTCCGAGCTACGCT 3300

RESULT 2

AAD11592

ID AAD11592 standard; DNA; 3300 BP.

XX AAD11592;

AC AAD11592;

DT 24-SEP-2001 (first entry)

DE Partial sequence of pMal-Ap3 and ZFPap3 DNA.

XX Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;

KW modulation; plant technology; agriculture; Ap3; APETALA3; ds.
 OS Unidentified.

XX Key Location/Qualifiers
 XX CDS 2719..3270
 FT /tag= a
 FT /product= "ZFPap3 protein"
 FT /transl_except= (pos:3124..3126, aa:Ser)
 FT /transl_except= (pos:3127..3129, aa:Ser)
 FT /transl_except= (pos:3136..3138, aa:Val)
 FT /note= "CDS does not include start and stop codon"
 FT /partial
 FT primer_bind 2740..2790
 FT /tag= b
 FT /bound_moiety= "F1-f2 primer"
 FT primer_bind 2770..2850
 FT /tag= c
 FT /bound_moiety= "F1-f1 primer"
 FT primer_bind 2824..2889
 FT /tag= d
 FT /bound_moiety= "F2-b primer"
 FT primer_bind 2867..2940
 FT /tag= e
 FT /bound_moiety= "F2-f primer"
 FT primer_bind 2916..2973
 FT /tag= f
 FT /bound_moiety= "F3-b1 primer"
 FT primer_bind 2953..3021
 FT /tag= g
 FT /bound_moiety= "F3-b2 primer"
 FT primer_bind 2992..3042
 FT /tag= h
 FT /bound_moiety= "F4-f2 primer"
 FT primer_bind 3022..3102
 FT /tag= i
 FT /bound_moiety= "F4-f1 primer"
 FT primer_bind 3076..3141
 FT /tag= j
 FT /bound_moiety= "F5-b primer"
 FT primer_bind 3119..3192
 FT /tag= k
 FT /bound_moiety= "F5-f primer"
 FT primer_bind 3168..3225
 FT /tag= l
 FT /bound_moiety= "F6-b1 primer"
 FT primer_bind 3205..3273
 FT /tag= m
 FT /bound_moiety= "F6-b2 primer"

XX WO200152620-A2.

XX 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US001817.

XX 21-JAN-2000; 2000US-017748P.

XX 21-JUL-2000; 2000US-00620897.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA AGRIC DISCOVERY INC.

XX Barbas CF, Stege JT, Guan X, Dalmia B;

XX WPI; 2001-465325/50.

XX P-PSDB; AAE06005.

XX New zinc finger proteins, useful for modulating or regulating gene expression and metabolic pathways in plants, e.g. for treating in the plant cells a disorder that is associated with abnormal expression of the target gene.

XX Example 4; Page 148-149; 156pp; English.

XX The patent discloses methods and compositions to modulate the expression of a target gene in plant cells. The method involves providing plant cells with a zinc finger protein (ZFP) which is capable of specifically binding to a target nucleotide sequence or its complementary strand within a target gene and allowing the ZFP binding to the target nucleotide sequence, where the expression of the target gene in the plant cells is modulated. The ZFP and fusions of the ZFP proteins are useful for modulating or regulating gene expression and metabolic pathways in plants. The ZFP, fusion proteins and methods are useful in plant and agricultural technology. The method is useful particularly for treating a disorder in the plant cells, where the disorder is associated with abnormal expression of the target gene. The present DNA sequence is the partial sequence of pMal-Ap3 (APETALA3) and ZFPap3 DNA

XX Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;

Query Match 98.0%; Score 3232.8; DB 4; Length 3300;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 3258; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy	1	CCGACACCATCGAATGTCGCAAAACCTTTCGCGTATGGCATGATACGCCCGGAAGAGA	60
Db	1	CCGACACCATCGAATGTCGCAAAACCTTTCGCGTATGGCATGATACGCCCGGAAGAGA	60
Qy	61	GTCAATTCAAGGTGTCGAATGTGAACCAAGTAAAGTTATACGATGTCGAGAGTATCCG	120
Db	61	GTCAATTCAAGGTGTCGAATGTGAACCAAGTAAAGTTATACGATGTCGAGAGTATCCG	120
Qy	121	GTGTCCTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCCAGCGTTCGCGAAAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCGTTCGCGAAAA	180
Qy	181	CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATACATTCACACCGGTGGCAC	240
Db	181	CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATACATTCACACCGGTGGCAC	240
Qy	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC	300
Qy	301	ACGCGCGTGCAGAAATTTGTCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCGTGCAGAAATTTGTCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Qy	361	TGTTGTTGTCGATGTTAGAACGAAGCGGCTCGAAGCGGCTGTAAGCGGCGGTGCACAATC	420
Db	361	TGTTGTTGTCGATGTTAGAACGAAGCGGCTCGAAGCGGCTGTAAGCGGCGGTGCACAATC	420
Qy	421	TTCTCGCGCAACCGTCAAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA	480
Db	421	TTCTCGCGCAACCGTCAAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA	480
Qy	481	TTGCTGTGGNAGCTGCTGCACATAATTTTCGCGCGTTATTTCTTGATGTCTTGACAGA	540
Db	481	TTGCTGTGGNAGCTGCTGCACATAATTTTCGCGCGTTATTTCTTGATGTCTTGACAGA	540
Qy	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC	600
Qy	601	TGTTGCGATTGGGTCAACAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCGTCTCGG	660
Db	601	TGTTGCGATTGGGTCAACAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCGTCTCGG	660
Qy	661	CGGTCCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCATCAAAATTCAGCCCATAG	720
Db	661	CGGTCCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCATCAAAATTCAGCCCATAG	720
Qy	721	CGGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
Db	721	CGGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA	780


```
Db 2941 AAGTCTTTCAGCGAGTCCAGCAACCTGGTGGCCACCAACGTACTACACCGGGGAGAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCAGCAGCAGTGGCTCTTGGTTAGACAC 3060
Qy 3061 CAGCGTACCCACACCGGTGAAACCGGTATTAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAACCGGTATTAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Qy 3121 CGCAGCGATAACCTGGTGGCCATCAAGCGCACTCATCTGGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTGGAGCGCCATCAAGCGCACTCATCTGGGAGAGCCATACAAATGT 3180
Qy 3181 CCAGAAATGCGCAAGTCTTTCTCAACTTCAGGCCATTTGGTCCGTCACCAACGTACTCAC 3240
Db 3181 CCAGAAATGCGCAAGTCTTTCTCAACTTCAGGCACTTTGGTCCGTCACCAACGTACTCAC 3240
Qy 3241 ACCGGTAAAAAACTAGTGGCGCAGCGCGCCGCGTACCGTACGAGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCGCAGCGCGCCGCGTACCGTACGAGTTCCGGACTACGCT 3300

RESULT 3
AAD11588
ID AAD11588 standard; DNA; 3300 BP.
XX
AC
XX
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m1 and ZFPm1 DNA.
XX
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key
FT Location/Qualifiers
FT 2719..3270
FT CDS
FT /*tag= a
FT /product= "ZFPm1 protein"
FT /transl_except= (pos:2785..2787, aa:Asp)
FT /transl_except= (pos:2788..2790, aa:Pro)
FT /transl_except= (pos:2791..2793, aa:Gly)
FT /transl_except= (pos:2872..2874, aa:Arg)
FT /transl_except= (pos:2875..2877, aa:Ala)
FT /transl_except= (pos:2878..2880, aa:His)
FT /transl_except= (pos:2884..2886, aa:Glu)
FT /transl_except= (pos:2853..2855, aa:Gln)
FT /transl_except= (pos:2959..2961, aa:Ser)
FT /transl_except= (pos:3121..3123, aa:Arg)
FT /transl_except= (pos:3124..3126, aa:Ser)
FT /transl_except= (pos:3127..3129, aa:Asp)
FT /transl_except= (pos:3130..3132, aa:Asn)
FT /transl_except= (pos:3136..3138, aa:Val)
FT /transl_except= (pos:3139..3141, aa:Arg)
FT /transl_except= (pos:3205..3207, aa:Gln)
FT /transl_except= (pos:3208..3210, aa:Ala)
FT /transl_except= (pos:3211..3213, aa:Gly)
FT /transl_except= (pos:3214..3216, aa:His)
FT /transl_except= (pos:3220..3222, aa:Ala)
FT /transl_except= (pos:3223..3225, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /*tag= d
```

```
FT primer_bind
FT /bound_moiety= "F2-b primer"
FT 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"
XX
XX WO200152620-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US001817.
XX
XX 21-JAN-2000; 2000US-0177468P.
XX
XX 21-JUL-2000; 2000US-00620897.
XX
XX (SCRI ) SCRIPPS RES INST.
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CF, Stege JT, Guan X, Dalmia B;
XX WPI; 2001-465325/50.
XX P-PSDB; AAE06000.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX
XX Example 4; Page 138-139; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX partial sequence of pMal-m1 and ZFPm1 DNA
XX
XX Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
SQ
```

```
Query Match 97.4%; Score 3215.2; DB 4; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
```

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAGAGA 60
QY 61 GTCAATTCCAGGGTGGTGAATGTGA AACCCAGTAAACGTTATACGATGTCGACAGTATGCCG 120
Db 61 GTCAATTCCAGGGTGGTGAATGTGA AACCCAGTAAACGTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGA ACCAGCCAGCCAGCCGTTTCTGCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGA ACCAGCCAGCCAGCCGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTA CATTCCCAACCGGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTA CATTCCCAACCGGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGGCGCTCGCAATTTCTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCGCAGCG 360
Db 301 ACGGCGCTCGCAATTTCTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCGCAGCG 360
QY 361 TGGTGGTTCGATGGTAGAAGCGGCGCTCGAAGCGCTGTAAAGCGGCGGTGCAACAATC 420
Db 361 TGGTGGTTCGATGGTAGAAGCGGCGCTCGAAGCGCTGTAAAGCGGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAAC TATCCGCTGATCACCAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAAC TATCCGCTGATCACCAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCCTGCAATAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCAATAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGATTTGGTCCACAGCAAAATCGCGCTGTTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGATTTGGTCCACAGCAAAATCGCGCTGTTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTC AACAACCAATGCNAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTC AACAACCAATGCNAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGATGCTGTTGCCA ACGATCGATGGCGTGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTCGATGCTGTTGCCA ACGATCGATGGCGTGGGCGCAA 840
QY 841 TGGCGGCCATTAACCGAGTCCGGGCTGCGGTTGGTGC GGAATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTAACCGAGTCCGGGCTGCGGTTGGTGC GGAATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTATGTTATATCCGCGCTTAA CCAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTATGTTATATCCGCGCTTAA CCAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGCGTGGACCGCTTGCTGCAACTCTC TAGGCCACAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGCGTGGACCGCTTGCTGCAACTCTC TAGGCCACAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGTGCCGCTCTCACTGGTGAAA AAAAAACCAACCCCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTGCCGCTCTCACTGGTGAAA AAAAAACCAACCCCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCAACGACAGGTTT 1140

Db 1081 CGCAAAACCGCCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCAACGACAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTTCAGAGCTTATCATGCACTGCACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTTCAGAGCTTATCATGCACTGCACGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACAGCCATCCGAAAGCTGTGTATGGCTGTCAGGTCGTAATACTCACTGCATAATTCG 1320
Db 1261 TCAGGACAGCCATCCGAAAGCTGTGTATGGCTGTCAGGTCGTAATACTCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAAATTTTTTTCGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAAATTTTTTTCGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATTAATCATCGGCTCGTATATGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATTAATCATCGGCTCGTATATGTGGA 1440
QY 1441 ATTCTGAGCGGATAACAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCACGA 1500
Db 1441 ATTCTGAGCGGATAACAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGCTGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGCTGTAATCTGG 1560
QY 1561 ATTTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAATCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAATCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGTTGAGCATCCCGATAACTGGAAGAGAAATCCCAACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACGTTGAGCATCCCGATAACTGGAAGAGAAATCCCAACAGTT 1680
QY 1681 CGGCAACTGCGGATGGCCCTTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGGATGGCCCTTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAAATCAACCCGGAAGAGGTTCCAGGAACAAGCTGAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAAATCAACCCGGAAGAGGTTCCAGGAACAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTTACAACGGCAAGCTGATTGCTTACC CGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTTACAACGGCAAGCTGATTGCTTACC CGATCGCTGT 1860
QY 1861 GAAAGCTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCGGCCAAAAACCTGGGAA 1920
Db 1861 GAAAGCTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCGGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAAAACAACAATGAGATGACAGACCGGATTC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAAAACAACAATGAGATGACAGACCGGATTC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220

Db 2161 TCCATCGCAGAGCTGCCCTTTTAAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGGTG 2220
Qy 2221 GCATGCTCAACATCGACACAGCAAGTGAATTTATGTTAAACGGTACTCGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGTGAATTTATGTTAAACGGTACTCGCGACCTTC 2280
Qy 2281 AAGGTCAACCATCAAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTTAAACGCCGCCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTCGTTGGCGTCTGAGCGCAGGTATTTAAACGCCGCCAGT 2340
Qy 2341 CGGAACAAAGAGCTGGCAAAAGAGTCTCGAAATATCTGCTACTGATGAAGTCTG 2400
Db 2341 CGGAACAAAGAGCTGGCAAAAGAGTCTCGAAATATCTGCTACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAAGAAACGCCCAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAAGAAACGCCCAAGGTGAATCATG 2520
Qy 2521 CGGAACATCCGCGAGATGTCGCTTTCTGTTATGCGGTATGCGGTATGCGGTATCAAGCC 2580
Db 2521 CGGAACATCCGCGAGATGTCGCTTTCTGTTATGCGGTATGCGGTATGCGGTATCAAGCC 2580
Qy 2581 GCCAGCGTCTGACATGTCGATGAAGCCTCGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCTCGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACAAACAATTAACAATAACAACCTCGGGATCGAGGAAGATTCAGAAATTC 2700
Db 2641 AACACAAACAATTAACAATAACAACCTCGGGATCGAGGAAGATTCAGAAATTC 2700
Qy 2701 GGATCCTCTTCTGTCGCGCAGCGCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCCTCTTCTGTCGCGCAGCGCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGTAAAGTCTTCTCTCAGAGCTCTCAGCTGGTGGCCACAGCTACCCAC 2820
Db 2761 CCGGAATGTGTAAAGTCTTCTCTCAGAGCTCTCAGCTGGTGGCCACAGCTACCCAC 2820
Qy 2821 ACGGGTGAACACCGTATAATGCCAGAGTGGCGCAATCTTTTAGCCAGCAGCAGC 2880
Db 2821 ACGGGTGAACACCGTATAATGCCAGAGTGGCGCAATCTTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Qy 2941 AAGTCTTTCTCGGTCTGCAATCTCGTCCGCGCACCACCAAGTACTCACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCGGTCTGCAATCTCGTCCGCGCACCACCAAGTACTCACCGGGGAGAAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCTCAGAGCTCTACCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCAGCGAGCGGATACCTGGTGGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
Qy 3121 CGAGCGATACCTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGGCCACCTGGCGAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Qy 3181 CCAGAAATGTGCAAGTCTTTCTCAACTTCAGGCCATTTGTCGTACCAACGCTACTCAC 3240
Db 3181 CCAGAAATGTGCAAGTCTTTCTCTCGGTCTGACAAATCTCTGCGGCAACGCTACTCAC 3240
Qy 3241 ACCGGTAAAAAATAGTGGCCAGCGCGGCAAGTACCCGTACGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGCGCGGCGCGCAGTACCCGTACGAGTTCGGGACTACGCT 3300

RESULT 4
AAD11589
ID AAD11589 standard; DNA; 3300 BP.
XX
AC AAD11589;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m2 and ZFPm2 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
XX modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
CDS 2719..3270
FT /*tag= a
FT /product= "ZFPm2 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"

WO200152620-A2.

26-JUL-2001.

19-JAN-2001; 2001WO-US001817.

21-JAN-2000; 2000US-0177468P.

21-JUL-2000; 2000US-00620897.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA AGRIC DISCOVERY INC.

Barbas CF, Stege JT, Guan X, Dalmia B;

	XX		WPI; 2001-465325/50.
	DR		P-ESDB; AAE06002.
	PT	New zinc finger proteins, useful for modulating or regulating gene expression and metabolic pathways in plants, e.g. for treating in the plant cells a disorder that is associated with abnormal expression of the target gene.	
	XX		Example 4; Page 140-142; 156pp; English.
	XX	The patent discloses methods and compositions to modulate the expression of a target gene in plant cells. The method involves providing plant cells with a zinc finger protein (ZFP) which is capable of specifically binding to a target nucleotide sequence or its complementary strand within a target gene and allowing the ZFP binding to the target nucleotide sequence, where the expression of the target gene in the plant cells is modulated. The ZFP and fusions of the ZFP proteins are useful for modulating or regulating gene expression and metabolic pathways in plants. The ZFP, fusion proteins and methods are useful in plant and agricultural technology. The method is useful particularly for treating a disorder in the plant cells, where the disorder is associated with abnormal expression of the target gene. The present DNA sequence is the partial sequence of pMal-m2 and ZFm2 DNA	
	XX		Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
	SQ		
		Query Match	97.4%; Score 3215.2; DB 4; Length 3300;
		Best Local Similarity	98.4%; Pred. No. 0;
		Matches 3247; Conservative	0; Mismatches 53; Indels 0; Gaps 0;
Oy	1	CCGACACCATCGAATGTGC AAAACCTTTTCGGTGATGCCATGATAGC CC CGGAAGA GA	
Db	1	CCGACACCATCGAATGTGTCAAACCTTTTCGGTGATGCCATGATAGC CC CGGAAGA GA	
Oy	61	GTCAA TTCAGGTGTGAATTGTGA AACCAACGTAACGTTTATACGATGTCGAGAGTATGCCG	120
Db	61	GTCAA TTCAGGTGTGTGAATGTGA AACCAACGTAACGTTTATACGATGTCGAGAGTATGCCG	120
Oy	121	TGTGTCCTTATCAG ACCGTTTTCCCGGTGGTGAACCGAGCCAGCACGTTCTGCGAAA	180
Db	121	TGTGTCCTTATCAG ACCGTTTTCCCGGTGGTGAACCGAGCCAGCACGTTCTGCGAAA	180
Oy	181	CGCGGAAAAAGTGAAAGCGGCGATGGCGAGCTGAATTACA TTCCAACCCGGTGGCAC	240
Db	181	CGCGGAAAAAGTGAAAGCGGCGATGGCGAGCTGAATTACA TTCCAACCCGGTGGCAC	240
Oy	241	AACAA CTGGCGGGCAAAACAGTCTGTTGATTTGGG GTGCCACTTCAGTCTGGCCCTGC	300
Db	241	AACAA CTGGCGGGCAAAACAGTCTGTTGATTTGGG GTGCCACTTCAGTCTGGCCCTGC	300
Oy	301	ACGGCCGTCCCAANTTGTCCGGCGATAAATCTCGCCGCGATCAATCGGGTGCCAGCG	360
Db	301	ACGGCCGTCCGCAANTTGTCCGGCGGATAAATCTCGCCGCGATCAATCGGGTGCCAGCG	360
Oy	361	TGGTGGTTCGATGTTAGAA CGAAGCGGCGTTCGAAGCCTGTAAAGCGCGGTGCACAATC	420
Db	361	TGGTGGTTCGATGTTAGAA CGAAGCGGCGTTCGAAGCCTGTAAAGCGCGGTGCACAATC	420
Oy	421	TTCTCGCCAAACGCTCAGTGGGCTGATCATTA ACTATCCGCTGGATGACCAAGGATGCCA	480
Db	421	TTCTCGCCAAACGCTCAGTGGGCTGATCATTA ACTATCCGCTGGATGACCAAGGATGCCA	480
Oy	481	TTCGTTGGAGAGCTGCCTGCAC TAATGTTCCGGCGTTATTTCTTGATGTCTGTGAC CAG	540
Db	481	TTCGTTGGAGAGCTGCCTGCAC TAATGTTCCGGCGTTATTTCTTGATGTCTCTTGAC CAG	540
Oy	541	CACCCATCAACAGTATTTTCTCCC ATGAAGACGGTAGCGGAC TGCGGCTGGAGCATC	600
Db	541	CACCCATCAACAGTATTTTCTCCC ATGAAGACGGTAGCGGAC TGCGGCTGGAGCATC	600
Oy	601	TGGTCGATTTGGGTCAACAGCAAATTCGCGCTGTTAGCGGGCCCATTAAGTTTGTCTCGG	660

601	Db	TG	GT	CG	CA	TT	TGG	GT	CA	CG	CA	AA	AT	CG	CG	TG	TT	TAG	CG	GG	CC	CA	TT	AA	GT	T	GT	T	GT	CT	CG	660				
661	Qy	CG	GT	CT	GC	CT	TG	CG	TG	GT	CG	CA	TAA	A	T	AT	CT	C	A	CT	CG	CA	A	T	T	C	A	A	T	T	C	A	CG	720		
661	Db	CG	GT	CT	GC	CT	TG	CG	TG	GT	CG	CA	TAA	A	T	AT	CT	C	A	CT	CG	CA	A	T	T	C	A	A	T	T	C	A	CG	720		
721	Qy	CG	AA	CG	GA	AG	CG	A	CT	GA	GT	GC	CA	T	GT	C	CG	TT	T	CA	A	CA	A	A	C	CA	T	GA	A	A	T	GT	CT	GA	780	
721	Db	CG	AA	CG	GA	AG	CG	A	CT	GA	GT	GC	CA	T	GT	C	CG	TT	T	CA	A	CA	A	A	C	CA	T	GA	A	A	T	GT	CT	GA	780	
781	Qy	AT	G	AG	G	CA	T	CG	T	CC	CA	C	T	CG	AT	CT	GC	T	GC	T	CC	CA	AG	A	T	CA	GA	T	CA	GA	T	CG	CG	CA	840	
781	Db	AT	G	AG	G	CA	T	CG	T	CC	CA	C	T	CG	AT	CT	GC	T	GC	T	CC	CA	AG	A	T	CA	GA	T	CA	GA	T	CG	CG	CA	840	
841	Qy	TG	CG	CG	CA	T	T	A	C	CG	AG	T	CC	GG	CT	GC	CG	T	GC	T	GC	GA	T	A	T	CT	CG	T	AG	T	GG	GA	T	AC	900	
841	Db	TG	CG	CG	CA	T	T	A	C	CG	AG	T	CC	GG	CT	GC	CG	T	GC	T	GC	GA	T	A	T	CT	CG	T	AG	T	GG	GA	T	AC	900	
901	Qy	AC	GA	T	AC	CG	AA	GA	CA	GT	CA	T	AT	T	A	T	AT	T	AT	T	CC	CG	CG	T	T	AA	CA	CA	A	CA	GA	AT	T	T	C	960
901	Db	AC	GA	T	AC	CG	AA	GA	CA	GT	CA	T	AT	T	A	T	AT	T	AT	T	CC	CG	CG	T	T	AA	CA	CA	A	CA	GA	AT	T	T	C	960
961	Qy	GC	T	GT	CT	GG	GG	CA	A	CC	AG	CG	T	GA	CG	CT	T	GC	CA	CT	CT	C	AG	G	CG	CA	CG	CG	CG	GT	GA	1020				
961	Db	GC	T	GT	CT	GG	GG	CA	A	CC	AG	CG	T	GA	CG	CT	T	GC	CA	CT	CT	C	AG	G	CG	CA	CG	CG	CG	GT	GA	1020				
1021	Qy	AG	GC	AA	T	CA	GT	T	GC	CC	GT	CT	C	A	CT	T	GC	T	GC	AA	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1080	
1021	Db	AG	GC	AA	T	CA	GT	T	GC	CC	GT	CT	C	A	CT	T	GC	T	GC	AA	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1080	
1081	Qy	CG	AA	AC	CG	CG	CT	CT	CC	CG	CG	T	T	GC	CA	T	TA	T	TA	T	GC	AG	CT	GC	CA	CG	CA	CG	GT	T	T	T	T	T	1140	
1081	Db	CG	AA	AC	CG	CG	CT	CT	CC	CG	CG	T	T	GC	CA	T	TA	T	TA	T	GC	AG	CT	GC	CA	CG	CA	CG	GT	T	T	T	T	T	1140	
1141	Qy	CC	GA	CT	GA	AA	CG	GG	CA	GT	GA	CG	CA	A	CG	CA	A	T	TA	T	GT	GA	GT	TA	GT	TA	GT	TA	GT	TA	GT	TA	GT	TA	1200	
1141	Db	CC	GA	CT	GA	AA	CG	GG	CA	GT	GA	CG	CA	A	CG	CA	A	T	TA	T	GT	GA	GT	TA	GT	TA	GT	TA	GT	TA	GT	TA	GT	TA	1200	
1201	Qy	GC	CA	A	T	T	C	A	T	G	T	T	T																							

```
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGCAAAAGCGTTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGCAAAAGCGTTTCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGCTGATTAACACGCGAAGCTGATTCCTTACCCGATCCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGCTGATTAACACGCGAAGCTGATTCCTTACCCGATCCTGTT 1860
QY 1861 GAAGCGTTATCCTCTGATTTATACCAAGATCTGCTGCGGAAACCCGCAAAACCTTGGGAA 1920
Db 1861 GAAGCGTTATCCTCTGATTTATACCAAGATCTGCTGCGGAAACCCGCAAAACCTTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGACGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGACGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCAACCTGGCGCTGATGTTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCAACCTGGCGCTGATGTTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATTAAGCGTGGCGGCAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATTAAGCGTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGCTGATTAAGAACTTAAACCAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGCTGATTAAGAACTTAAACCAACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTAAATAAGCGCAACAGCGATGACCATCAACGGCGCGGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTAAATAAGCGCAACAGCGATGACCATCAACGGCGCGGTGG 2220
QY 2221 GCATGGTCAACATCGACACACGCAAAAGTGAATTAAGGTAAACGCTACCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGCAAAAGTGAATTAAGGTAAACGCTACCGACCTTC 2280
QY 2281 AAGGTCAACCATCAAAACCGTTGCTGGCTGCTGAGCGAGGTATTAACGCCCGCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTGCTGGCTGCTGAGCGAGGTATTAACGCCCGCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGGTTCTCGAACAATCTCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGGTTCTCGAACAATCTCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTACGAGGAAG 2460
QY 2461 TTGGGAAAGATCCACGTATTGCGCGCAACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTATTGCGCGCAACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCTGCGGTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCTGCGGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACACACAAATACAAATACAACTCCGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACACAAATACAAATACAACTCCGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTCGCGCGCGGCTCTGAGCGCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGCGCGCGGCTCTGAGCGCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAGTCTTCTGAGCGCAGAGAGCTCTCTGCTGCGCCACACGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAGTCTTCTCTCTCAGAGCTCTCAGCTGCTCAGCTGCTGCGCCACGCGTACCCAC 2820
```

RESULT 5

```
AAD11590
ID AAD11590 standard; DNA; 3300 BP.
XX
AC AAD11590;
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m3 and ZFPm3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key
FT CDS
FT 2719..3270
FT /*tag= a
FT /product= "ZFPm3 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT 2740..3290
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"
FT 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT primer_bind
```



```
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTCCGCTCAAGCGGCACATCCCGTTCTGGATATGTTTTTGGCCGACATCATAAACGGTT 1380
Db 1321 TGTCCGCTCAAGCGGCACATCCCGTTCTGGATATGTTTTTGGCCGACATCATAAACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGCGA 1440
Db 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGCGA 1440
QY 1441 ATTGTAGCGGATACAATTTTCAACAGAAACAGCCAGTCGTTAGGTGTTTTTCAAGA 1500
Db 1441 ATTGTAGCGGATACAATTTTCAACAGAAACAGCCAGTCGTTAGGTGTTTTTCAAGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAACCTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAACCTGAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTTAACGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAACGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCTACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCTACAGGTT 1680
QY 1681 GCGGCAACTGCGGATGCGCCCTGACATTAATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGCGGATGCGCCCTGACATTAATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGTGGCTGAATCAACCCGGACAAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGTGGCTGAATCAACCCGGACAAAAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTAAACGGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTAAACGGCAAGCTGATGCTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATPAAAGAACTGAAACGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATPAAAGAACTGAAACGAAAGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCCGCTGATTCGCTGACGGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCCGCTGATTCGCTGACGGGGGTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTTAAAGACGTGGCGTGGATACGCTGGCGGAAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTTAAAGACGTGGCGTGGATACGCTGGCGGAAAAGCG 2100
QY 2101 GGTCTGACCTTCCCTGGTTGACCTGATTTAAACAAAAACAATGAATGAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCCCTGGTTGACCTGATTTAAACAAAAACAATGAATGAGACACCGGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGAGCGAACAACAGGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGAGCGAACAACAGGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAACATCGACACCGCAAAAGTGAATTTATGGTGAACGCTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACCGCAAAAGTGAATTTATGGTGAACGCTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGCGTGGCTGCTGAGCGGAGGTATTAAACGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGCGTGGCTGCTGAGCGGAGGTATTAAACGCGCCAGT 2340
QY 2341 CCGAAACAAAGACTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
```

```
Db 2341 CCGAAACAAAGACTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCCGCTGGGTGCGTAGCGCTGAAAGTCTTTACAGGAAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCCGCTGGGTGCGTAGCGCTGAAAGTCTTTACAGGAAAG 2460
QY 2461 TTGGGGAAGATCCACGTTATTTGCCGCCACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGGAAGATCCACGTTATTTGCCGCCACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
QY 2701 GGATCTCTTCTCTGTCGCGCCAGCGCGCTCTCGAGCCCGGGGAGAAAGCCCTATGTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGCGCCAGCGCGCTCTCGAGCCCGGGGAGAAAGCCCTATGTTGT 2760
QY 2761 CCGGAATGTGGTAAAGTCTCTTTCAGCAGAGCAGCTCCTGTCGCGCCACCGAGCGTACCAC 2820
Db 2761 CCGGAATGTGGTAAAGTCTCTTTCAGCAGTCTCTGGCCACCTGGTTGCCACCGAGCTACCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTCGGGCAAAATCTTTTAGCCAGAGCAGCAGC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTCGGGCAAAATCTTTTAGCCAGAGCAGCAGC 2880
QY 2881 CTGGTCGCCATCAACGCACTCATCTGCGGAGAAGCCATACAAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTCGCCATCAACGCACTCATCTGCGGAGAAGCCATACAAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTTCAAGTGTGCTGATCTTTCAGAGGACCAACGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTTCAAGTGTGCTGATCTTTCAGAGGACCAACGTACTCACACCGGGGAGAAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTCTCTCAGAGCTCTCACCTGTCGCGCAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTCTCAGAGAGCAGCTCCCTGTCGCGCAC 3060
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTCCGCAAAATCTTTTAGT 3120
Db 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTCCGCAAAATCTTTTAGT 3120
QY 3121 CGCAGCGATAACTGTCGCGCATCAACGCACTCATCTGCGGAGAAGCCATACAAATGT 3180
Db 3121 GACTGCGCGACCTGCTCGCCATCAACGCACTCATCTGCGGAGAAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCAAATTCAGGCGCAATTTGGTCCGTCAACCAACGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTGTCGCGGACCAACGTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCAGTACCCGTAACGCTTCGCGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCAGTACCCGTAACGCTTCGCGACTACGCT 3300
```

RESULT 6

AAH25681

ID AAH25681 standard; DNA; 8101 BP.

XX

AC AAH25681;

XX

DT 05-SEP-2001 (first entry)

XX

DE Nucleotide sequence of human Cdc25C phosphatase DNA.

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
 KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200144467-A2.
 XX
 PD 21-JUN-2001.
 XX
 XX 13-DEC-2000; 2000WO-FR003496.
 PF
 XX 14-DEC-1999; 99FR-00015722.
 PR
 FR 30-MAY-2000; 2000FR-00006883.
 PR
 XX 21-SEP-2000; 2000FR-00012008.
 XX
 PA (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
 XX
 PI Goubin-Gramatica F, Ducommun B, Prevost G;
 XX
 DR WPI; 2001-398152/42.
 XX
 PT New fusion protein of maltose-binding protein and Cdc phosphatase, useful
 PT for identifying phosphatase modulators for regulating the cell cycle.
 XX
 PS Claim 7; Page 9-13; 56pp; French.
 XX
 CC The present sequence encodes human Cdc25C phosphatase, and was expressed
 CC in Escherichia coli. It is used to make fusion proteins with Escherichia
 CC coli maltose binding protein (MBP). The specification describes fusion
 CC proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
 CC are involved in activation of cyclin-dependent kinases (CDK) that control
 CC cell division. The fusion proteins are used to identify modulators of the
 CC specified human Cdc phosphatases, potentially useful for regulation of
 CC cell division. They are also for studying physiological or
 CC physiopathological activities of Cdc phosphatase
 XX
 SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;
 Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CGACACCATCGAATGGTGCAGAAACCTTTCGCGGTATGCGATGATAGCCGCCGGAAGAGA 60
 Db 1 CGACACCATCGAATGGTGCAGAAACCTTTCGCGGTATGCGATGATAGCCGCCGGAAGAGA 60
 Qy 61 GTCATTCAGGGTGTGATGTGAACACGATTAACGTTATACGATGTCGAGAGTATGCCG 120
 Db 61 GTCATTCAGGGTGTGATGTGAACACGATTAACGTTATACGATGTCGAGAGTATGCCG 120
 Qy 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGCTTTCGCGAANA 180
 Db 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGCTTTCGCGAANA 180
 Qy 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
 Db 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
 Qy 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
 Db 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
 Qy 301 ACGCCCGTCCGAAATTTGCGGCGGATTAATCTCGCGCGCATCACTGGGTGCCAGCG 360
 Db 301 ACGCCCGTCCGAAATTTGCGGCGGATTAATCTCGCGCGCATCACTGGGTGCCAGCG 360
 Qy 361 TGGTGGTGTGATGTAGAACGAGCGCGTGCAGAGCTGTAAAGCGGCGGTGCACAATC 420
 Db 361 TGGTGGTGTGATGTAGAACGAGCGCGTGCAGAGCTGTAAAGCGGCGGTGCACAATC 420
 Qy 421 TTCTCGCGAAACGCGTCACTAGTGGGCTGATCACTAACTATCCCGTGTGATGACCAAGGATGCCA 480

Db 421 TTCTCGCGAAACGCGTCACTAGTGGGCTGATCACTAATATCCCGTGTGATGACCAAGATGCCA 480
 Qy 481 TTGCTGTGGAAGAGTCCCTGCGACTAATGTTCCGGCGTTTATTTCTGATGCTCTGTGACCA 540
 Db 481 TTGCTGTGGAAGAGTCCCTGCGACTAATGTTCCGGCGTTTATTTCTGATGCTCTGTGACCA 540
 Qy 541 CACCATCAACAGTATTATTTCTCCCATGAAGAGGTACCGGACTGGCGCTGGAGCATC 600
 Db 541 CACCATCAACAGTATTATTTCTCCCATGAAGAGGTACCGGACTGGCGCTGGAGCATC 600
 Qy 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCCAATTAAGTTCTGTCTCG 660
 Db 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCCAATTAAGTTCTGTCTCG 660
 Qy 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
 Db 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
 Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCAACAAACCAATGCAATGCTGA 780
 Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCAACAAACCAATGCAATGCTGA 780
 Qy 781 ATGAGGCGATGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGGCTGGGCGCAA 840
 Db 781 ATGAGGCGATGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGGCTGGGCGCAA 840
 Qy 841 TGCAGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATAG 900
 Db 841 TGCAGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATAG 900
 Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACACCAATCAACAGGATTTTC 960
 Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACACCAATCAACAGGATTTTC 960
 Qy 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
 Db 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
 Qy 1021 AGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAAACCAACCTGGCGCCCAATA 1080
 Db 1021 AGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAAACCAACCTGGCGCCCAATA 1080
 Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGGTTT 1140
 Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGGTTT 1140
 Qy 1141 CCGACTGGAAAGCGGCGAGTGAGCGCAACCAATTAATGAGTGTAGTCACTCATTTAG 1200
 Db 1141 CCGACTGGAAAGCGGCGAGTGAGCGCAACCAATTAATGAGTGTAGTCACTCATTTAG 1200
 Qy 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAACGGTGCACCAATGTTCTGGCG 1260
 Db 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAACGGTGCACCAATGTTCTGGCG 1260
 Qy 1261 TCAGGACCCATTCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1320
 Db 1261 TCAGGACCCATTCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1320
 Qy 1321 TGTGCTCAAGCGGCGACTCCCGTTCCTGGATTAATGTTTTCGCGCGAGCATCAACAGGTT 1380
 Db 1321 TGTGCTCAAGCGGCGACTCCCGTTCCTGGATTAATGTTTTCGCGCGAGCATCAACAGGTT 1380
 Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCACAATTAATCATCGGCTCGTATATGTTGTA 1440
 Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCACAATTAATCATCGGCTCGTATATGTTGTA 1440
 Qy 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAAACAGCAGTCCGTTTAGGTTTTCAGCA 1500
 Db 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAAACAGCAGTCCGTTTAGGTTTTCAGCA 1500
 Qy 1501 GCATTTCAACAAAGGACCATAGTTATGAACCTGAAGAGGTAACACTGGTATCTCG 1560

Db 1501 GCACCTTCACCAACAAGGACCATAGATTATGAAAAATCGAAGAGGTAAATCTGGTATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATTAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATTAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTACACGGTTGAGCATCCGGATAAAGTGGAGAGAAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTACACGGTTGAGCATCCGGATAAAGTGGAGAGAAAATTCACACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCCGGACAAAGGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCCGGACAAAGGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTACGTTTACACCGCAAGCTGATTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTACACCGCAAGCTGATTTACCCGATCGCTGTT 1860
Qy 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTGCCGAACCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTGCCGAACCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
Qy 1981 CTGCAAGAACCTGATTAACCTGGCGCTGATTTGCTGTCAGCGGGGTTATGCTTTCAAG 2040
Db 1981 CTGCAAGAACCTGATTAACCTGGCGCTGATTTGCTGTCAGCGGGGTTATGCTTTCAAG 2040
Qy 2041 TATGAAGAACGCAAGTACGACATTAAGACCTGGCGCTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAGAACGCAAGTACGACATTAAGACCTGGCGCTGGATACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAACCAACACACATGAATGACACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAACCAACACACATGAATGACACACCGGATTAC 2160
Qy 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAACTCGACACCAAGTGAATTAATGCTGTAAGCTGCTGCGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACCAAGTGAATTAATGCTGTAAGCTGCTGCGACCTTC 2280
Qy 2281 AAGGGTCAACATCCAAACCGTTGTTGGGCTGCTGAGCGCAGGTAATTAACGGCCAGT 2340
Db 2281 AAGGGTCAACATCCAAACCGTTGTTGGGCTGCTGAGCGCAGGTAATTAACGGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGCAGGTATTAACGGCCAGT 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGCAGGTATTAACGGCCAGT 2460
Qy 2461 TTGGCGAAGATCCACGTATTTGCCGCCACCATGGAACCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTATTTGCCGCCACCATGGAACCCGAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGGCTTTCTGCTGATGCGGTGCTGCTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGGCTTTCTGCTGATGCGGTGCTGCTGCGGTGATCAAGCC 2580
Qy 2581 GCCAGCGGTGCTGAGACTGCTGATGAAGCCCTGAAAGACCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGAGACTGCTGATGAAGCCCTGAAAGACCGCAGACTAATTCGAGCTCG 2640

Qy 2641 AACCAACAACAATAACAATAACAACAACCTCGGGATCGAGGAAGGATTTCAAGATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACAACCTCGGGATCGAGGAAGGATTTCAAGATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 7
ABL53238
ID ABL53238 standard; DNA; 8101 BP.
XX ABL53238;
AC ABL53238;
XX
DT 17-JUN-2002 (first entry)
XX
Nucleotide sequence of pMAL-Hs Cdc25C.
XX
Cdc25C phosphatase; amine; Cdc25 phosphatase; spontaneous alopecia;
KW alopecia; proliferative disease; parasitic disease; viral infection;
KW neurodegenerative disease; myopathy; angiogenesis; psoriasis; restenosis;
KW vitamin K; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 1528..2677
FT /*tag= b
FT /note= "encodes maltose binding protein (MBP)"
FT CDS 2713..4134
FT /*tag= a
FT /note= "ORF of human Cdc25C"
XX
PN FR2812198-A1.
XX
PD 01-FEB-2002.
XX
PF 28-JUL-2000; 2000FR-00009900.
XX
PR 28-JUL-2000; 2000FR-00009900.
XX
PA (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
PI Prevost G, Brezak PMC, Galcera CMO, Thureau C, Goubin GF;
PI Ducommun B;
XX
DR WPI; 2002-166066/22.
XX
XX
PT cdc25 Phosphatase inhibiting medicaments, useful e.g. for treating tumor
PT diseases, viral infections, neurodegenerative disease or alopecia,
PT containing new or known aryl-substituted amines.
XX
PS
XX
XX
CC The present sequence represents pMAL-HsCdc25C, a plasmid encoding human
CC Cdc25C phosphatase. The specification describes the use of aryl-
CC substituted secondary or tertiary amines for the production of
CC medicaments for inhibiting Cdc25 phosphatases, especially Cdc25C
CC phosphatases. The amines of the invention are used for treating
CC spontaneous alopecia or alopecia induced by exogenous products or
CC radiation. They are also used for treating tumoral or non-tumoral
CC proliferative diseases, parasitic diseases, viral infections,
CC neurodegenerative diseases or myopathy. Examples of non-tumoral
CC proliferative diseases are e.g. angiogenesis, psoriasis or restenosis.
CC The amines are also useful for inhibiting the proliferation of
CC microorganisms (especially yeasts) and in the treatment of all disorders
CC conventionally treated with vitamin K or its derivatives
XX
SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match 82.0%; Score 2705.8; DB 6; Length 8101; Best Local Similarity 99.9%; Pred. No. 0; Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
Qy	1	CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATATAGAGCCCGGAAGAGA	60						
Db	1	CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATATAGAGCCCGGAAGAGA	60						
Qy	61	GTCAATTACAGGGTGTGAATGTGAACACAGTAACCTTTATACGATCTCGAGATATGCG	120						
Db	61	GTCAATTACAGGGTGTGAATGTGAACACAGTAACCTTTATACGATCTCGAGATATGCG	120						
Qy	121	GTGTCCTCTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCCACGTTTCTGCGAAA	180						
Db	121	GTGTCCTCTATCAGACCGTTTCCCGGTGGTGAACAGGCGCAGCCACGTTTCTGCGAAA	180						
Qy	181	CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTAATCCCAACCGCTGGGCAC	240						
Db	181	CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTAATCCCAACCGCTGGGCAC	240						
Qy	241	AACAACTGGCGGCAAAAAGTGTGCTGATTTGGCGTTGCCACTTCCAGTCTGGCCCTGC	300						
Db	241	AACAACTGGCGGCAAAAAGTGTGCTGATTTGGCGTTGCCACTTCCAGTCTGGCCCTGC	300						
Qy	301	ACGCGCCCTCGCAAAATTTGTCGGGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGG	360						
Db	301	ACGCGCCCTCGCAAAATTTGTCGGGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGG	360						
Qy	361	TGGTGTGTCGATGTGAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGCTGCACAATC	420						
Db	361	TGGTGTGTCGATGTGAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGCTGCACAATC	420						
Qy	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA	480						
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA	480						
Qy	481	TTGCTGTGGAGCTGCCCTGCATTAATGTTCCGGCGTTATTTCTTGTATGTTCTGACCCAGA	540						
Db	481	TTGCTGTGGAGCTGCCCTGCATTAATGTTCCGGCGTTATTTCTTGTATGTTCTGACCCAGA	540						
Qy	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGGGCTGGGCGTGGAGCATC	600						
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGGGCTGGGCGTGGAGCATC	600						
Qy	601	TGGTCCGATTTGGGTCAACAGCAAAATCGCGCTTTAGCGGGGCCATTAAGTTCTGTCTCG	660						
Db	601	TGGTCCGATTTGGGTCAACAGCAAAATCGCGCTTTAGCGGGGCCATTAAGTTCTGTCTCG	660						
Qy	661	CGCGTCTGGCTGTGGCTGGCATTAATATCTCACTCGGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGCGTCTGGCTGTGGCTGGCATTAATATCTCACTCGGCAATCAAAATTCAGCCGATAG	720						
Qy	721	CGGAACGGGAAGGCGACTGGAGTGCATCTCCGGTTTTCACAAACCATGCAAAATGCTGA	780						
Db	721	CGGAACGGGAAGGCGACTGGAGTGCATCTCCGGTTTTCACAAACCATGCAAAATGCTGA	780						
Qy	781	ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCGGCTGGGCGGCA	840						
Db	781	ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCGGCTGGGCGGCA	840						
Qy	841	TGCGGCGCATTAACGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG	900						
Db	841	TGCGGCGCATTAACGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG	900						
Qy	901	ACGATACCGAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAAAAGGATTTTC	960						
Db	901	ACGATACCGAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAAAAGGATTTTC	960						
Qy	961	GCCTCTGGGGCAAAACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020						
Db	961	GCCTCTGGGGCAAAACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020						
Qy	1021	AGGCAATACGCTGTTGCCCGTCTCACTGGTGA AAAAACAACATGAATGCAACCCGATTC	1080						

Db	1021	AGGCAATACGCTGTTGCCCGTCTCACTGGTGA AAAAACAACCCCTGGCGCCCAATA	1080						
Qy	1081	CGCAAAACCGCTCTCCCCCGCGCTTGGCCGATTCATTAATGACGTGGCAGCACAGGTTT	1140						
Db	1081	CGCAAAACCGCTCTCCCCCGCGCTTGGCCGATTCATTAATGACGTGGCAGCACAGGTTT	1140						
Qy	1141	CCGACCTGGAAGCGGCGAGTGAGGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200						
Db	1141	CCGACCTGGAAGCGGCGAGTGAGGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200						
Qy	1201	GCACAAATTTCTCATGTTTGACAGCTTATCATGACGTGACAGGTGACCAATGCTTCTGGCG	1260						
Db	1201	GCACAAATTTCTCATGTTTGACAGCTTATCATGACGTGACAGGTGACCAATGCTTCTGGCG	1260						
Qy	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCG	1320						
Db	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCG	1320						
Qy	1321	TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT	1380						
Db	1321	TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT	1380						
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGA	1440						
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGA	1440						
Qy	1441	ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCGTTTAGGTGTTTTCAAGA	1500						
Db	1441	ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCGTTTAGGTGTTTTCAAGA	1500						
Qy	1501	GCATTTCAACCAAGGACCATAGATTATGA AAAACTGAAGAGGTAAACTGTTAATCTGG	1560						
Db	1501	GCATTTCAACCAAGGACCATAGATTATGA AAAACTGAAGAGGTAAACTGTTAATCTGG	1560						
Qy	1561	ATTAACCGCGATAAGGCTATAACGGTCTCCGCTGAAGTCGCTAGAAATTCGAGAAAGAT	1620						
Db	1561	ATTAACCGCGATAAGGCTATAACGGTCTCCGCTGAAGTCGCTAGAAATTCGAGAAAGAT	1620						
Qy	1621	ACCGGAATTAAGTACCGTTTGAAGCATCCGGATTAACCTGGAAGAGAAAATTCGCCAGGTT	1680						
Db	1621	ACCGGAATTAAGTACCGTTTGAAGCATCCGGATTAACCTGGAAGAGAAAATTCGCCAGGTT	1680						
Qy	1681	CGGCAACTGCGGATGGCCCTTGACATATCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Db	1681	CGGCAACTGCGGATGGCCCTTGACATATCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Qy	1741	GCTCAATCTGGCGCTTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAGCTGAT	1800						
Db	1741	GCTCAATCTGGCGCTTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAGCTGAT	1800						
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAACCGCAAGCTGATTGCTTACCGATCGCTGTT	1860						
Db	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAACCGCAAGCTGATTGCTTACCGATCGCTGTT	1860						
Qy	1861	GAAAGCTTATCGCTGATTTATAACAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Db	1861	GAAAGCTTATCGCTGATTTATAACAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Qy	1921	GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980						
Db	1921	GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980						
Qy	1981	CTGCAAGAAACCGTACTTCCCTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG	2040						
Db	1981	CTGCAAGAAACCGTACTTCCCTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG	2040						
Qy	2041	TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAACG	2100						
Db	2041	TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAACG	2100						
Qy	2101	GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAACAACATGAATGCAACCCGATTC	2160						

```
Db 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTTAAAAACAACATGAATGACGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAAGCTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACACGAAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGAAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCTTGGCGTCTGAGCGAGGTATTAAGCGCCCGAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTTGGCGTCTGAGCGAGGTATTAAGCGCCCGAGT 2340
QY 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGGAAAGATCCACGATTGCGCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGATTGCGCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGCGCGCAGACTAAATCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGCGCGCAGACTAAATCGAGTCTG 2640
QY 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGATTTCAGAAATC 2700
Db 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGATTTCAGAAATC 2700
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
AD023608
ID AD023608 standard; DNA; 7259 BP.
AC
AC AD023608;
XX
XX 01-JUL-2004 (first entry)
XX
XX DNA encoding MBP-Toxop30del18(82-294aa) fusion protein.
XX
XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; Gene.
XX
XX Toxoplasma gondii.
OS Synthetic.
XX
XX US2004067239-A1.
XX
XX 08-APR-2004.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX (MAIN/) MAINE G T.
PA (FATE/) FATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
```

PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;

XX WPI; 2004-304563/28.

DR P-PSDB; ADO23609.

XX Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of Igm or Igg
PT antibodies to Toxoplasma gondii.

XX Example 2; Fig 22; 114pp; English.

PS The invention relates to a purified P30 antigen (I) chosen from 3 fully
XX defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of Igm
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the Igm antibodies
CC with a composition comprising (I) and detecting the presence of (I)/Igm
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.

XX Sequence 7259 BP; 1870 A; 1822 C; 1926 G; 1641 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 CCACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTTCAGGCTGGTGAATGTGAAACCAAGTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTTCAGGCTGGTGAATGTGAAACCAAGTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCGCCAGCCAGCTTTCTGCGAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCGCCAGCCAGCTTTCTGCGAAA 180
QY 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATACATTCACACCGCGTGGCAC 240
Db 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATACATTCACATTCACACCGCGTGGCAC 240
QY 241 AACAACTGCGCGGCAACACAGTCTGCTGATTTGGCGTGGCCACTCCAGTCTCGCCCTGC 300
Db 241 AACAACTGCGCGGCAACACAGTCTGCTGATTTGGCGTGGCCACTCCAGTCTCGCCCTGC 300
QY 301 ACGCCGCTCGCAAAATGTCGCGCGAATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCGCTCGCAAAATGTCGCGCGAATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGATGGTAGAACGCGGTGGAAGCGCTGTAAGCGCGGTGTAAGCGCGGTGCAATC 420
Db 361 TGGTGTGTGATGGTAGAACGCGGTGGAAGCGCTGTAAGCGCGGTGTAAGCGCGGTGCAATC 420
QY 421 TTCTCGCGCAACCGCTGAGTGGGCTGATCATTAATCTATCCGCTGGATGACACGAGTGC 480
Db 421 TTCTCGCGCAACCGCTGAGTGGGCTGATCATTAATCTATCCGCTGGATGACACGAGTGC 480
QY 481 TTGCTGTGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACAGA 540
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGAGCGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGAGCGTACGCGACTGGCGGTGGAGCATC 600
```

QY 601 TGGTCGCATTCGGGTACCAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTCTCTCGG 660
Db 601 TGGTCGCATTCGGGTACCAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTCTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGGATGAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGGATGAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACCAATGCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACCAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGCGGCTGGGCGAA 840
Db 781 ATGAGGGCATGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGCGGCTGGGCGAA 840
QY 841 TGGCGCCATTAACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCCATTAACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACGAGATTTC 960
Db 901 ACGATACCGAAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAACGAGATTTC 960
QY 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAAGAAACACACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAAGAAACACACCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTCAGACGTTTATCATGACCTGCACGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTCAGACGTTTATCATGACCTGCACGGTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGCAGGTTCGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGCAGGTTCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGCATCATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGCATCATCAACGGTT 1380
QY 1381 CTGGCAATATCTCAAAATGAGCTGTGCAATTAATCATCGGCTCGTATATGTCGA 1440
Db 1381 CTGGCAATATCTCAAAATGAGCTGTGCAATTAATCATCGGCTCGTATATGTCGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTCAACAGGAACAGCCAGTCGGTTAGGTGTTTCAAGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTCAACAGGAACAGCCAGTCGGTTAGGTGTTTCAAGA 1500
QY 1501 GCACCTTCCAAACAAAGGACCATAGATTATGAAAACTGAAGAAAGGTAATCTGGTATCTGG 1560
Db 1501 GCACCTTCCAAACAAAGGACCATAGCATATGAAAAATCGAAGAAAGGTAATCTGGTATCTGG 1560
QY 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740

RESULT 9
ADO23613
ID ADO23613 standard; DNA; 7322 BP.
XX AC ADO23613;

Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGGAAGCTGATGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGGAAGCTGATGCTTACCCGATCGCTGTT 1860
QY 1861 GAACGCTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAACGCTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGAATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGAATAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATGATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATGATGACAGACCGGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAAGAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACAGCAAAAGTGAATTTATGCTGTAACGCTGACCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAAAGTGAATTTATGCTGTAACGCTGACCGACCTTC 2280
QY 2281 AAGGCTCAACCATCAAAACCGTTGCTGCTGCGCAGGTATTAACCGCCCACT 2340
Db 2281 AAGGCTCAACCATCAAAACCGTTGCTGCTGCGCAGGTATTAACCGCCCACT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAATATCTGCTGACGTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAATATCTGCTGACGTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGCAAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGCAAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
QY 2461 TTGCGAAAAGATCCACGTTATTTGCCGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGAAAAGATCCACGTTATTTGCCGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACAATCCCGCAGATGTCGCGTTTCTGTTGATGCGGTGCGTATCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACAATCCCGCAGATGTCGCGTTTCTGTTGATGCGGTGCGTATCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTCGTCAGACTGTCGATGAAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCGTCAGACTGTCGATGAAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACCAACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

QY	1	CGGACCAATCGAATGGTGCAGAAACCTTTCGGGTATGGCATGATGAGCCCGGAGAGA	60
DB	1	CGGACCAATCGAATGGTGCAGAAACCTTTCGGGTATGGCATGATGAGCCCGGAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTATACGATGTCGACAGTATGCG	120
DB	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTATACGATGTCGACAGTATGCG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTATACGATGTCGACAGTATGCG	180
DB	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTATACGATGTCGACAGTATGCG	180
QY	181	CGCGGAAAGTGAAGCGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC	240
DB	181	CGCGGAAAGTGAAGCGCGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCAAAACAGTCTGTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGGCGGCAAAACAGTCTGTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTCCGAATTTCTCGCGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGC	360
DB	301	ACGCGCGTCCGAATTTCTCGCGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGC	360
QY	361	TGGTGGTCTCGATGTAAGCGCGCTGAAAGCTGTAATTAATCTCGCGCGATCAACTGGGTGCCAGC	420
DB	361	TGGTGGTCTCGATGTAAGCGCGCTGAAAGCTGTAATTAATCTCGCGCGATCAACTGGGTGCCAGC	420
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAGC	480
DB	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAGC	480
QY	481	TTGCTGTGGAGCTTCGCTGCACTAATGTTCCGGGCTTATTTCTTGATGTTCTCGACAGA	540
DB	481	TTGCTGTGGAGCTTCGCTGCACTAATGTTCCGGGCTTATTTCTTGATGTTCTCGACAGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGCGTGGAGCATC	600
QY	601	TGGTGGATTTGGGTACACAGCAATTCGGTGTAGCGGGCCCAATTAAGTTCTGCTCGG	660
DB	601	TGGTGGATTTGGGTACACAGCAATTCGGTGTAGCGGGCCCAATTAAGTTCTGCTCGG	660
QY	661	CGGCTCTGGCTGGCTGGCTGCAATTAATCTCACTCGCAATCAATTCAGCCGATAG	720
DB	661	CGGCTCTGGCTGGCTGGCTGCAATTAATCTCACTCGCAATCAATTCAGCCGATAG	720
QY	721	CGGAAACGGGAGGCGATGGAGTGGCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780
DB	721	CGGAAACGGGAGGCGATGGAGTGGCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCAATCGTTCCCACTCGCATGCTGGTGGCAACGATGAGTGGCGCTGGCGCAA	840
DB	781	ATGAGGCAATCGTTCCCACTCGCATGCTGGTGGCAACGATGAGTGGCGCTGGCGCAA	840
QY	841	TGGCGGCAATTCAGAGTCCGGGCTGGCGTGGTGGCAATCTCGTATGTTGGATACG	900
DB	841	TGGCGGCAATTCAGAGTCCGGGCTGGCGTGGTGGCAATCTCGTATGTTGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCATCAACAGGATTTTC	960

QY	961	GCCTGCTGGGCAAAACAGCGGTGACCCGCTTGCTGCAACTCTCTCAGGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGCAAAACAGCGGTGACCCGCTTGCTGCAACTCTCTCAGGCCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTCCTGCTCTCAGTGTGAAAGAAACACACCTGGCGGCCAATA	1080
DB	1021	AGGCAATCAGCTGTCCTGCTCTCAGTGTGAAAGAAACACACCTGGCGGCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGGTGGCGATTAATTAATGACAGTGGCACACAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGGTGGCGATTAATTAATGACAGTGGCACACAGGTTT	1140
QY	1141	CCGCACTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG	1200
DB	1141	CCGCACTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGTGGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGTGGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCAATCGGAAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	1320
DB	1261	TCAGGACGCAATCGGAAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	1320
QY	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAACCGTT	1380
DB	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAACCGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTGTATAATGTTGGA	1440
DB	1381	CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTGTATAATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCGAGTCCGTTTGTGTTTCAACA	1500
DB	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCGAGTCCGTTTGTGTTTCAACA	1500
QY	1501	GCATTTCAACAAACAGGACCATAGATTTGAAATCTGAAAGTAACTGAAAGTAACTG	1560
DB	1501	GCATTTCAACAAACAGGACCATAGATTTGAAATCTGAAAGTAACTGAAAGTAACTG	1560
QY	1561	ATTAAACGGGATTAAGGCTATTAACCGTCTCGCTGAAAGTAACTGAAAGTAACTG	1620
DB	1561	ATTAAACGGGATTAAGGCTATTAACCGTCTCGCTGAAAGTAACTGAAAGTAACTG	1620
QY	1621	ACCGGAATTAAGTCAACCGTGTGAGCATCGGATTAACCTGAAAGTAACTGAAAGT	1680
DB	1621	ACCGGAATTAAGTCAACCGTGTGAGCATCGGATTAACCTGAAAGTAACTGAAAGT	1680
QY	1681	GGCGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACAGCGCTTTGGTGGCTAC	1740
DB	1681	GGCGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACAGCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAGCTGTAT	1800
DB	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGGCGTGTGATTAACCGGACAGCTGATTTACCGGATCGCTGTT	1860
DB	1801	CGGTTTACCTGGGATGGCGTGTGATTAACCGGACAGCTGATTTACCGGATCGCTGTT	1860
QY	1861	GAAGCGTATTCGCTGATTTTATAACAAAGATCTGCTGCCAAACCGGCAAAACCTGGGAA	1920
DB	1861	GAAGCGTATTCGCTGATTTTATAACAAAGATCTGCTGCCAAACCGGCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAAACCGTACTTTCACCTGGCGCTGATGTTGCTGACAGGCGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAAACCGTACTTTCACCTGGCGCTGATGTTGCTGACAGGCGGTTATCGGTTCAAG	2040

541	Db	CACCCATCAACAGTAATATTTTCTCCCATGAAGACGGTACGCGACTGGGGCTGGAGCATC	600
601	Qy	TGGTCGCATTTGGGTACCAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG	660
601	Db	TGGTCGCATTTGGGTACCAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG	660
661	Qy	CGCGTCTGGCTGTGGCTGGCTGGCATAAATATCTCACTCGGCAATCAAAATTCAGCGGATAG	720
661	Db	CGCGTCTGGCTGTGGCTGGCTGGCATAAATATCTCACTCGGCAATCAAAATTCAGCGGATAG	720
721	Qy	CGGAACGGGAAGCGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780
721	Db	CGGAACGGGAAGCGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780
781	Qy	ATGAGGGCATGTTTCCCATCTGCGATGCTGGTGTGCCAAGATCAAGATGCGCTGGGCGCAA	840
781	Db	ATGAGGGCATGTTTCCCATCTGCGATGCTGGTGTGCCAAGATCAAGATGCGCTGGGCGCAA	840
841	Qy	TGCGGCCATTACCGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGTAGTGGGATACG	900
841	Db	TGCGGCCATTACCGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGTAGTGGGATACG	900
901	Qy	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAAAACGAGATTTTC	960
901	Db	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAAAACGAGATTTTC	960
961	Qy	GCCTGTGGGGCAAAACACGCTGGACCGCTTGCTGCACTCTCTCAGGCGCAGCGGTGA	1020
961	Db	GCCTGTGGGGCAAAACACGCTGGACCGCTTGCTGCACTCTCTCAGGCGCAGCGGTGA	1020
1021	Qy	AGGGCAATCAGCTGTGTGCCCTCTCACTGGTGAAGAAACACACCTCGCGCCCAATA	1080
1021	Db	AGGGCAATCAGCTGTGTGCCCTCTCACTGGTGAAGAAACACACCTCGCGCCCAATA	1080
1081	Qy	CGCAAAACCGCTCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCAGCAGGTTT	1140
1081	Db	CGCAAAACCGCTCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCAGCAGGTTT	1140
1141	Qy	CCCGACTGGAAAGCGGGCAGTGAGCGCAACCGCAATTAATGTAGTTAGTTCACTCATTAG	1200
1141	Db	CCCGACTGGAAAGCGGGCAGTGAGCGCAACCGCAATTAATGTAGTTAGTTCACTCATTAG	1200
1201	Qy	GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCACGGTGCAACAAATGCTTCGCG	1260
1201	Db	GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCACGGTGCAACAAATGCTTCGCG	1260
1261	Qy	TCAGCAGCCATCGAAGCTGTGATGGCTGTCAGGTGCTGAATCACTGCATAAATTCG	1320
1261	Db	TCAGCAGCCATCGAAGCTGTGATGGCTGTCAGGTGCTGAATCACTGCATAAATTCG	1320
1321	Qy	TGTCGCTCAAGGCGCATCTCCGTTTCTGGATAATGTTTTTTCGCGCGACATCATACGCTT	1380
1321	Db	TGTCGCTCAAGGCGCATCTCCGTTTCTGGATAATGTTTTTTCGCGCGACATCATACGCTT	1380
1381	Qy	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
1381	Db	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
1441	Qy	ATTGTAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTTCAGCA	1500
1441	Db	ATTGTAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTTCAGCA	1500
1501	Qy	GCATTCACCAACAGGACCATAGATTATGAAACTGAAGGTAACCTGGTAAATCTCG	1560
1501	Db	GCATTCACCAACAGGACCATAGATTATGAAACTGAAGGTAACCTGGTAAATCTCG	1560
1561	Qy	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGGTGAAGAAATTCGAGAAGAT	1620
1561	Db	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGGTGAAGAAATTCGAGAAGAT	1620
1621	Qy	ACCGGAATTAAGTCAACCGTTTGACATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680
1621	Db	ACCGGAATTAAGTCAACCGTTTGACATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680

Qy	1691	CGCGAACTGCGGATGCGCCCTGACATTAATCTTCTGGGCAACGACACCGCTTTGGTGGCTAC	1740
Db	1691	CGCGAACTGCGGATGCGCCCTGACATTAATCTTCTGGGCAACGACACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACACAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCGGTAGTGTTAAACGGCAAGCTGATGCTTACCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCGGTAGTGTTAAACGGCAAGCTGATGCTTACCCGATCGCTGTT	1860
Qy	1861	GAACGGTTATCGCTGATTTATAACAAGATCTCTCGCGAACCGCGCAAAAACCTGGAA	1920
Db	1861	GAACGGTTATCGCTGATTTATAACAAGATCTCTCGCGAACCGCGCAAAAACCTGGAA	1920
Qy	1921	GAGATCCGGGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGGGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTTCACTCGCGCCTGATTTGCTGTCGACGGGGGTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTTCACTCGCGCCTGATTTGCTGTCGACGGGGGTATGCGTTCAAG	2040
Qy	2041	TATGAAAAACGGCAAGTAGCAATTAAGACGCTGGGCGTGGATTAACGCTGCGCGAAAGCG	2100
Db	2041	TATGAAAAACGGCAAGTAGCAATTAAGACGCTGGGCGTGGATTAACGCTGCGCGAAAGCG	2100
Qy	2101	GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGAATGACAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGAATGACAGACACCGATTAC	2160
Qy	2161	TCCATCCGAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG	2220
Db	2161	TCCATCCGAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG	2220
Qy	2221	GCATGGTCCAAACATCGACACACGAAAGTGAAATTAATGTTGTGTAACGCTACTGCCGACCTTC	2280
Db	2221	GCATGGTCCAAACATCGACACACGAAAGTGAAATTAATGTTGTGTAACGCTACTGCCGACCTTC	2280
Qy	2281	AAGGTCACACATCCAAACCGTTTGTGGCGTCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Db	2281	AAGGTCACACATCCAAACCGTTTGTGGCGTCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Qy	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
Qy	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTACGAGGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTACGAGGGAAGAG	2460
Qy	2461	TTGGCAAAAGATCCACGTATTGCGCGCACCATGGGAAAAACGCCAGAAAGGTGAAATCATG	2520
Db	2461	TTGGCAAAAGATCCACGTATTGCGCGCACCATGGGAAAAACGCCAGAAAGGTGAAATCATG	2520
Qy	2521	CCGAAACATCCCGAGATGTCGCGTTTCTGGTATGCGGTGCGTACTCGCGTGATCAACGCC	2580
Db	2521	CCGAAACATCCCGAGATGTCGCGTTTCTGGTATGCGGTGCGTACTCGCGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGTCAGACTGTGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGTCAGACTGTGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG	2640
Qy	2641	AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2641	AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC	2700

RESULT 12
ADO23639
ID ADO23639 standard; DNA; 7370 BP.

XX AD023639;
 AC
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE DNA encoding MBP-ToxoP30MIX1 fusion protein.
 XX
 KW P30 antigen; Toxo30del13C; Toxo30del2C; ToxoP30 MIX1;
 KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; ds; gene.
 XX
 OS Toxoplasma gondii.
 OS Synthetic.
 XX
 FN US2004067239-A1.
 XX
 PD 08-APR-2004.
 XX
 PF 02-OCT-2002; 2002US-00263153.
 XX
 PR 02-OCT-2002; 2002US-00263153.
 XX
 PA (MAIN/) MAINE G T.
 PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 XX
 FI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX
 XX MPI; 2004-304563/28.
 DR P-PSDB; AD023640.
 XX
 PT Novel purified polypeptide having sequence identity to amino acid
 PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX
 PS Example 5; Fig 31; 114pp; English.
 XX
 CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
 CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX
 SQ Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;
 Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CGGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
 DB 1 CGGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
 QY 61 GTCATTCAGGGTGGTGAATGTGAACAGTAACGTATACGATCTCCGAGATATGCGG 120
 DB 61 GTCATTCAGGGTGGTGAATGTGAACAGTAACGTATACGATCTCCGAGATATGCGG 120
 QY 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAA 180
 DB 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAA 180
 QY 181 CGCGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGC 240

DB 181 CGCGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
 QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGGTTTGGCCACCTCCAGTCTGGCCCTGC 300
 DB 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGGTTTGGCCACCTCCAGTCTGGCCCTGC 300
 QY 301 ACGGCGGTTCGCAAAATTTGTCGGGGGATTAATCTCGGCGCGATCAATCGGTGGTGCAGG 360
 DB 301 ACGGCGGTTCGCAAAATTTGTCGGGGGATTAATCTCGGCGCGATCAATCGGTGGTGCAGG 360
 QY 361 TGGTGGTGTGATGATGAAGAGCGGCTGGAAGCTGTAAAGCGGCGGTGCAAAATC 420
 DB 361 TGGTGGTGTGATGATGAAGAGCGGCTGGAAGCTGTAAAGCGGCGGTGCAAAATC 420
 QY 421 TTCTCGCGCAACGCGTCACTGCGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTATG 480
 DB 421 TTCTCGCGCAACGCGTCACTGCGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTATG 480
 QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGGTATTTCTTGATGTTCTGACACAGA 540
 DB 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGGTATTTCTTGATGTTCTGACACAGA 540
 QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGCTGGCGGTGGAGCATC 600
 DB 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGCTGGCGGTGGAGCATC 600
 QY 601 TGGTGCATTTGGGTCCACGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTCTGCTCGG 660
 DB 601 TGGTGCATTTGGGTCCACGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTCTGCTCGG 660
 QY 661 CGCGTCTGCGTCTGGCTGGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
 DB 661 CGCGTCTGCGTCTGGCTGGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
 QY 721 CGGAACGGGAAGGCGACGTGAGTGCATGTCGGTTCATCAACACCATGCAAAATGCTGA 780
 DB 721 CGGAACGGGAAGGCGACGTGAGTGCATGTCGGTTCATCAACACCATGCAAAATGCTGA 780
 QY 781 ATGAGGGCATGCTTCCCACTCGCATGCTGTTTGGCAACGATCAGATGGCGCTGGCGGCA 840
 DB 781 ATGAGGGCATGCTTCCCACTCGCATGCTGTTTGGCAACGATCAGATGGCGCTGGCGGCA 840
 QY 841 TGGCGGCCATTAACGAGTCCCGGCTGGCGGTGGGATATCTCGGTAGTGGGATAG 900
 DB 841 TGGCGGCCATTAACGAGTCCCGGCTGGCGGTGGGATATCTCGGTAGTGGGATAG 900
 QY 901 ACGATACCGAAGACGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC 960
 DB 901 ACGATACCGAAGACGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC 960
 QY 961 GCCTGCTGGGCAACACGAGCGTGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
 DB 961 GCCTGCTGGGCAACACGAGCGTGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
 QY 1021 AGGCAATCAGCTGTTGCCCGCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
 DB 1021 AGGCAATCAGCTGTTGCCCGCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
 QY 1081 CGCAACCGCTCTCTCCCGCGGTGGCGGATTCATTAATGCAAGCTGGCAGCAGAGTTT 1140
 DB 1081 CGCAACCGCTCTCTCCCGCGGTGGCGGATTCATTAATGCAAGCTGGCAGCAGAGTTT 1140
 QY 1141 CCCGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
 DB 1141 CCCGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
 QY 1201 GCACAATTTCTCATGTTTGCACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
 DB 1201 GCACAATTTCTCATGTTTGCACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
 QY 1261 TCAGGCAGCCATCGGAAGCTGTGTTATGGCTGTGAGGTGTAATCAATCACTGATTAATTCG 1320

CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (i) and detecting the presence of (i)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ

Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;
Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATCGCATATAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATCGCATATAGCGCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAACAGTTATACGATGTCGACAGATGCGG 120
Db 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAACAGTTATACGATGTCGACAGATGCGG 120
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCAGGCCACGTTTCTGCGAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCAGGCCACGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
QY 241 AACAACTCGCGGGGAAAAACAGTCTGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTCGCGGGGAAAAACAGTCTGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGGCGCTGCGCAATGTGCGGCGGATTAATCTCGCGCGGATCAAATGCGGTGCCAGCG 360
Db 301 ACGGCGCTGCGCAATGTGCGGCGGATTAATCTCGCGCGGATCAAATGCGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGTAAGCGGCGATGAACTATCTCGCGCGGATCAAATGCGGTGCCAGATC 420
Db 361 TGGTGGTTCGATGTAAGCGGCGGCTGAAAGCGGCTGAAAGCGGCGGTTGCAACATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGATCAGGAGTGC 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGATCAGGAGTGC 480
QY 481 TTGCTGTGGAAGCTCGCTGCAATAATGTTCGGCGGTTATTTCTGATGTCTTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTCGCTGCAATAATGTTCGGCGGTTATTTCTGATGTCTTGACCCAGA 540
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
QY 601 TGGTGGCATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db 601 TGGTGGCATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGGATCGTTCCCACTGCGATGCTGGTTGCCAAGCATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGGATCGTTCCCACTGCGATGCTGGTTGCCAAGCATCAGATGCGCTGGCGGCA 840
QY 841 TGCAGCGCATTAACCGAGTCCCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGCAGCGCATTAACCGAGTCCCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATATCCCGCGGTTAAACCAATCAAAACGAGATTTC 960

Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACATCAAAACGAGATTTC 960
QY 961 GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGTCCTCTCACTGCTGGAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTCCTCTCACTGCTGGAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAAATCAGCTGCGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAAATCAGCTGCGACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGAAGCTGTGGTATGCGTGTGAGGTCTGTAATCACTGCAATAATTCG 1320
Db 1261 TCAGGACGCAATCGAAGCTGTGGTATGCGTGTGAGGTCTGTAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTCTTGGATTAATGTTTTTTCGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTCTTGGATTAATGTTTTTTCGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTGCAACAATTAATCATCGGCTCGTAAATGCTGTA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTGCAACAATTAATCATCGGCTCGTAAATGCTGTA 1440
QY 1441 ATTGTGAGCGGATTAACAAATTTTTCACAGAAACAGCAGTCCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATTAACAAATTTTTCACAGAAACAGCAGTCCGTTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAATCTGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAATCTGTAATCTGG 1560
QY 1561 ATTTAAGCGCGATAAAGGCTATAACCGTCTCGCTGAGTTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGCGATAAAGGCTATAACCGTCTCGCTGAGTTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 CGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGCGGCAAGGCTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGCGGCAAGGCTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCGGAAACCCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCGGAAACCCGCCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGTAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGTAAGAGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATCGGTTCAAG 2040

```
Db 1981 CTGCAAGAACGTAATCTCACTGCGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAAGACGTCGGGTGATACGCTGCGGGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAAGACGTCGGGTGATACGCTGCGGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACACATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACACATGAATGACACACCGATTAC 2160
Qy 2161 TCCATCGCAGACGCTGCTTTAATAAGCGGAACGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGACGCTGCTTTAATAAGCGGAACGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGTCACCAATCGACACGACGATTAATGATGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGTCACCAATCGACACGACGATTAATGATGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGGTCACCAATCGACACGATTCGTTGGGTGCTGAGCGAGGATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCACCAATCGACACGATTCGTTGGGTGCTGAGCGAGGATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTAATAAGACAAACCGTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAGACAAACCGTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCACGATTCGCGCCACCAATGGAAGGAAAGCCGCAAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTCGCGCCACCAATGGAAGGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGATCAAGCC 2580
Qy 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGGAGACTAATTCAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGGAGACTAATTCAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700

RESULT 14
ID ADO23644 standard; DNA; 7370 BP.
AC ADO23644;
XX 01-JUL-2004 (first entry)
DE DNA encoding MBP-Toxop30MIX3 fusion protein.
XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX Toxoplasma gondii.
OS Synthetic.
XX US2004067239-A1.
XX 08-APR-2004.
XX 02-OCT-2002; 2002US-00263153.
XX 02-OCT-2002; 2002US-00263153.
XX (MAIN/) MAINE G T.
PA
```

```
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
XX (BLIE/) BLIESE T R.
PI Maine GT, Patel CB, Gineburg SR, Bliese TR;
XX WPI: 2004-304563/28.
DR P-PSDB; ADO23645.
XX Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX Example 5; Fig 33; 114pp; English.
XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATCGTGCACAAACCTTTCGGGTATGCGATGATGATGCGGGAAGAGA 60
Db 1 CCGACACCATCGAATCGTGCACAAACCTTTCGGGTATGCGATGATGATGCGGGAAGAGA 60
Qy 61 GTCATTCAGGGTGGTGAATGAAACCAAGTAACTGATATACGATGTCGAGAGTATGCCG 120
Db 61 GTCATTCAGGGTGGTGAATGTAACACCAAGTAACTGATATACGATGTCGAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTTGAAACAGCCAGCCAGCCAGCTTTCTGGGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTTGAAACAGCCAGCCAGCCAGCTTTCTGGGAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGCGCGGGAACAGTCTGTTGCTGATTTGGGCTTCCCACTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGCGGGAACAGTCTGTTGCTGATTTGGGCTTCCCACTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGCTGCGCAAAATTTGTCGGCGGATTAATATCTCGCGCGATCAACTGGTGGCAGCG 360
Db 301 ACGCGCGCTGCGCAAAATTTGTCGGCGGATTAATATCTCGCGCGATCAACTGGTGGCAGCG 360
Qy 361 TGGTGGTGTGATGTTAGAAACGAGCGCGTGAAGCTGTAAGCGCGGTGACCAATC 420
Db 361 TGGTGGTGTGATGTTAGAAACGAGCGCGTGAAGCTGTAAGCGCGGTGACCAATC 420
Qy 421 TTCTCGCGCAACGCGTCAAGTGGCTGATCAATTAATATCCGCTGGATGACGAGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCAAGTGGCTGATCAATTAATATCCGCTGGATGACGAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGGGTATTCTTTGATGTCCTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGGGTATTCTTTGATGTCCTGACACAGA 540
```

QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGACTCGCGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGACTCGCGCTGGAGCATC 600
QY 601 TGGTCGCATTTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGCGGCA 840
Db 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGCGGCA 840
QY 841 TGGCGGCCATTAACGAGTCCGGGCTGGCGGTTGGTGGCATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTAACGAGTCCGGGCTGGCGGTTGGTGGCATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAACCACCATCAACAGGATTTTC 960
Db 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAACCACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCCCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCCCTGGCGGCCAATA 1080
QY 1081 CGCAAAACGCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACGCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCAAGTGCACCAATGCTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCAAGTGCACCAATGCTCTGGCG 1260
QY 1261 TCAGGCAGCCATCGGAAGCTGTGTTATGCTGTGAGGTGCTGTAATCACTGCAATTTTCG 1320
Db 1261 TCAGGCAGCCATCGGAAGCTGTGTTATGCTGTGAGGTGCTGTAATCACTGCAATTTTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTTGAGCGGATTAACAAATTTTCAACAGAAACAGCCAGTCCGTTAGTGTTTTCAGA 1500
Db 1441 ATTTGAGCGGATTAACAAATTTTCAACAGAAACAGCCAGTCCGTTAGTGTTTTCAGA 1500
QY 1501 GCACCTTCAACAAAGGACCATAGCATATGAAAAATCGAAGAAAGGTAATCTGGTATCTGG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGCATATGAAAAATCGAAGAAAGGTAATCTGGTATCTGG 1560
QY 1561 ATTTAAGCGGATTAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAAGCGGATTAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGGAAGAGAAATTTCCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGGCAACTGCGCATGGCCCTTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGCATGGCCCTTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACTCGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCCTGTT 1860
Db 1801 CCGTTTACTCGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGCTGACGGGGGTTATCGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGCTGACGGGGGTTATCGCTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGCGCTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGCGCTGGATAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGCGAGACACCGATTAAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGCGAGACACCGATTAAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACACGCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACACGCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCAAACCATCAAAACCGTTGCTGGCGTGTGAGCGAGGTATTAACCGCCCGAGT 2340
Db 2281 AAGGTCAAACCATCAAAACCGTTGCTGGCGTGTGAGCGAGGTATTAACCGCCCGAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAANAACCTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAANAACCTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTTGCCGCACTATGGAACACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTTGCCGCACTATGGAACACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCG 2580
QY 2581 GCCAGCGGTCTGACAGTCTGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCTGACAGTCTGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGGAAGGATTTTCAGAAATTC 2700

AD023594
ID AD023594 standard; DNA; 7403 BP.
XX
AC ADO23594;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-ToxoP30del2(52-311aa) fusion protein.
XX
KW P30 antigen; Toxo30del3C; Toxo30del2C; ToxoP30 MIX1;
KW MBP-Toxo30del3C(52-300aa); MBP-Toxo30del4C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
WP1; 2004-304563/28.
DR P-PSDB; ADO23595.
DR
XX
XX Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del3C,
PT Toxo30del2C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
XX Example 2; Fig 11; 114pp; English.
XX
XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del3C, Toxo30del2C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del2C P30 antigen sequence, or comprising an amino acid sequence of
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del3C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del3C(52-300aa), MBP-Toxo30del4C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX
SQ Sequence 7403 BP; 1907 A; 1863 C; 1961 G; 1672 T; 0 U; 0 Other;
Query Match 81.5%; Score 2690.4; DB 12; Length 7403;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0
QY 1 CCGACACCATCGAATGGTGCAAAACCTTCGCGGTATGGCATATAGCCGCCGGAAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTCGCGGTATGGCATATAGCCGCCGGAAGA 60
QY 61 GTCAATTGAGGTGGTGAATGTGAAACACGATTAACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTGAGGTGGTGAATGTGAAACACGATTAACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACGCGCAGCTTCTCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACGCGCAGCTTCTCGAAAA 180

QY 1261 TCAGGACGCATCGGAAGCTGGGTATGGCTGTGCAGGTGTAATCACTGATTAATTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGGGTATGGCTGTGCAGGTGTAATCACTGATTAATTCG 1320
QY 1321 TGTCCCTCAAGCGGCACCTCCCGTTCCTGGATATGTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTCCCTCAAGCGGCACCTCCCGTTCCTGGATATGTTTTTGGCCGACATCATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTTGACAAATTAATCATCGGCTGTAATGTTGTGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTTGACAAATTAATCATCGGCTGTAATGTTGTGA 1440
QY 1441 ATTGTAGCGGATACAAATTTTACACAGCAACAGCCAGTCCGTTAGTGTGTTTCAG 1500
Db 1441 ATTGTAGCGGATACAAATTTTACACAGCAACAGCCAGTCCGTTAGTGTGTTTCAG 1500
QY 1501 GCACCTTCCACCAAGGACCATAGATTTATGAAAATGAAAGAGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCCACCAAGGACCATAGATTTATGAAAATGAAAGAGTAAACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATTAACGGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATTAACGGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTGGGCACAGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTGGGCACAGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGCGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGCGCAAGCTGATGCTTACCCGATCGCTGT 1860
QY 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTCCGCAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTCCGCAACCCGCCAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGACGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACAATCAATGAATCGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACAATCAATGAATCGACACCGGATTAC 2160
QY 2161 TCCATCGCAGAAAGCTGCCTTTAATAAGCGGAAACAGCGATCAACATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCCTTTAATAAGCGGAAACAGCGATCAACATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCCAACTCGACACCGCAAGTGAATTTATGGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACCGCAAGTGAATTTATGGTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCCTTCGAAAACCTATCTGCTGATGAAGGTCTG 2400

Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCCGCTGGGTGCGCTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCCGCTGGGTGCGCTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAAGATCCACGTAATTTGCCGCCACCATGGAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAAGATCCACGTAATTTGCCGCCACCATGGAAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTAATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTAATGCGGTGCTGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGCTGACACTGTCGATGAAGCCCTGAAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGCTGACACTGTCGATGAAGCCCTGAAAAGACGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAATTAACAATAACAACCTCGGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAATTAACAATAACAACCTCGGGATCGAGGAGGATTTTCAGAAATTC 2700

Search completed: August 1, 2005, 04:28:52
Job time : 1102.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-17
Perfect score: 3300
Sequence: 1 ccgacacacgaatgtgc.....acgacgttcgcgaactagct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818139359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	2700	81.8	7475	2	US-08-971-036-1	Sequence 1, Appli	
2	2700	81.8	7475	3	US-09-096-570-1	Sequence 1, Appli	
3	2700	81.8	7475	3	US-09-265-617B-1	Sequence 1, Appli	
C	4	1417.6	43.0	5926	3	US-09-027-169-3	Sequence 2, Appli
	5	1362	41.3	5201	4	US-09-640-882-2	Sequence 2, Appli
6	1362	41.3	5201	4	US-09-640-882-3	Sequence 3, Appli	
7	1204	36.5	4557	4	US-08-778-717-5	Sequence 5, Appli	
8	1200.8	36.4	3832	1	US-08-148-675A-2	Sequence 2, Appli	
C	9	1199.2	36.3	5248	3	US-08-487-283A-18	Sequence 18, Appl
C	10	1199.2	36.3	5248	5	PCT-US96-05611A-21	Sequence 21, Appl
C	11	1199.2	36.3	5312	4	US-10-263-103-35	Sequence 35, Appl
C	12	1199.2	36.3	5443	2	US-08-929-967-1	Sequence 1, Appli
C	13	1199.2	36.3	5502	4	US-09-702-705-785	Sequence 785, App
C	14	1199.2	36.3	5502	4	US-09-736-457-785	Sequence 785, App
C	15	1199.2	36.3	5502	4	US-09-614-124B-785	Sequence 785, App
C	16	1199.2	36.3	5502	4	US-09-671-325-785	Sequence 785, App
C	17	1199.2	36.3	5502	4	US-09-589-184-785	Sequence 785, App
C	18	1199.2	36.3	5502	4	US-09-658-824-785	Sequence 785, App
C	19	1199.2	36.3	5616	2	US-08-929-967-3	Sequence 3, Appli
C	20	1199.2	36.3	5873	4	US-09-695-437A-62	Sequence 62, Appl
C	21	1199.2	36.3	6353	4	US-09-702-705-784	Sequence 784, App
C	22	1199.2	36.3	6353	4	US-09-736-457-784	Sequence 784, App
C	23	1199.2	36.3	6353	4	US-09-614-124B-784	Sequence 784, App
C	24	1199.2	36.3	6353	4	US-09-671-325-784	Sequence 784, App
C	25	1199.2	36.3	6353	4	US-09-589-184-784	Sequence 784, App
C	26	1199.2	36.3	6353	4	US-09-658-824-784	Sequence 784, App
27	1199.2	36.3	6361	3	US-08-646-538-7	Sequence 7, Appli	

28	1199.2	36.3	6361	3	US-09-503-222-7	Sequence 7, Appli
C	29	1199.2	36.3	6727	2	Sequence 2, Appli
C	30	1199.2	36.3	6727	2	Sequence 2, Appli
C	31	1199.2	36.3	6799	2	Sequence 5, Appli
C	32	1199.2	36.3	6799	2	Sequence 5, Appli
C	33	1199.2	36.3	7659	3	Sequence 4, Appli
C	34	1199.2	36.3	7659	3	Sequence 213, App
C	35	1199.2	36.3	7676	3	Sequence 208, App
C	36	1199.2	36.3	7676	4	Sequence 213, App
C	37	1199.2	36.3	7676	4	Sequence 9, Appli
C	38	1199.2	36.3	8031	3	Sequence 254, App
C	39	1199.2	36.3	8031	4	Sequence 254, App
C	40	1199.2	36.3	8031	4	Sequence 254, App
C	41	1199.2	36.3	8031	4	Sequence 254, App
C	42	1199.2	36.3	8031	4	Sequence 254, App
C	43	1199.2	36.3	8031	4	Sequence 254, App
C	44	1199.2	36.3	8157	3	Sequence 1, Appli
C	45	1199.2	36.3	8501	3	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8% Score 2700; DB 2; Length 7475;

Best Local Similarity 100.0%; Pred. No. 0;				
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA	60	
Db	1	CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA	60	
QY	61	GTCAATTCAGGGTGGTGAATGTGA AACCCAGTAAAGTTATACGATGTGCACAGTATGCCG	120	
Db	61	GTCAATTCAGGGTGGTGAATGTGA AACCCAGTAAAGTTATACGATGTGCACAGTATGCCG	120	
QY	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTGA ACCAGGCAGCCACGTTTCTCGAAAA	180	
Db	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTGA ACCAGGCAGCCACGTTTCTCGAAAA	180	
QY	181	CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTACATCCCAACCGGTGCGAC	240	
Db	181	CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTACATCCCAACCGGTGCGAC	240	
QY	241	AACAACTGGCGGGCAAAACAGTCGTGTGTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC	300	
Db	241	AACAACTGGCGGGCAAAACAGTCGTGTGTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC	300	
QY	301	ACGCGCCGTGCGAAATTTGTCCGGCGGATTTAAATCTCGCGCGATCAACTGGGTGCCACGC	360	
Db	301	ACGCGCCGTGCGAAATTTGTCCGGCGGATTTAAATCTCGCGCGATCAACTGGGTGCCACGC	360	
QY	361	TGTTGGTGTGATGTTAGAACGAAGCGCGTGC GAAGCCTGTAAAGCGCGGTGCACATC	420	
Db	361	TGTTGGTGTGATGTTAGAACGAAGCGCGTGC GAAGCCTGTAAAGCGCGGTGCACATC	420	
QY	421	TTCTCGCGCAACCGTCCAGTGGGCTGATCAATTAATATCGCGTGGATGACACAGATGCCA	480	
Db	421	TTCTCGCGCAACCGTCCAGTGGGCTGATCAATTAATATCGCGTGGATGACACAGATGCCA	480	
QY	481	TTGCTGTGGAAGTCCCTGCACTAAATGTTCCGGCGTTATTTCTGTATGTTCTTGACGAGA	540	
Db	481	TTGCTGTGGAAGTCCCTGCACTAAATGTTCCGGCGTTATTTCTGTATGTTCTTGACGAGA	540	
QY	541	CACCCATCAACAGTATTTATTTTCTCCCATCAAGACGGTACCGCACTGGCGGTGGAGCATC	600	
Db	541	CACCCATCAACAGTATTTATTTTCTCCCATCAAGACGGTACCGCACTGGCGGTGGAGCATC	600	
QY	601	TGTTGCAATTTGGTTCACAGCAAAATCGCGTGTGTAGCGGGCCCATTAAGTTCGTCTCGG	660	
Db	601	TGTTGCAATTTGGTTCACAGCAAAATCGCGTGTGTAGCGGGCCCATTAAGTTCGTCTCGG	660	
QY	661	CGGCTCTGGCTCTGGCTGGTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
Db	661	CGGCTCTGGCTCTGGCTGGTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
QY	721	CGGAAACGGGAAGCGAGCTGGAGTGCCATGTC CGGTTTTCAACAAAACATGCAAAATCTGA	780	
Db	721	CGGAAACGGGAAGCGAGCTGGAGTGCCATGTC CGGTTTTCAACAAAACATGCAAAATCTGA	780	
QY	781	ATGAGGGCATCTGTTTCCACTGCGATGCTGGTGTGCCAACGATCAGATGGCGCTGGGCGCAA	840	
Db	781	ATGAGGGCATCTGTTCCCACTGCGATGCTGGTGTGCCAACGATCAGATGGCGCTGGGCGCAA	840	
QY	841	TGCGCGCCATTACCGAGTCCGGCTCGGCTGGTTCGGATATCTCGGTAGTGGGATACG	900	
Db	841	TGCGCGCCATTACCGAGTCCGGCTCGGCTGGTTCGGATATCTCGGTAGTGGGATACG	900	
QY	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGGTTAAACCACTCAAAACAGGATTTTC	960	
Db	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGGTTAAACCACTCAAAACAGGATTTTC	960	
QY	961	GCTTGTGGGGCAAAACAGGTGGAACCGTTGCTGCAACTCTCTCAGGCGCAGGGGTGA	1020	
Db	961	GCTTGTGGGGCAAAACAGGTGGAACCGTTGCTGCAACTCTCTCAGGCGCAGGGGTGA	1020	
QY	1021	AGGGCAATCAGCTGTTGGCGCTCACTGGTGAAGAAAAACCACTCTGGCGCCCAATA	1080	

Db	1021	AGGCAATCAGCTGTGTCGCCGTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA	1081
Qy	1081	CGCAAAACCGCTCTCCCCGGCGTGGCCGATTCATTAATGACGAGCTGGCACGACAGGTTT	1141
Db	1081	CGCAAAACCGCTCTCCCCGGCGTGGCCGATTCATTAATGACGAGCTGGCACGACAGGTTT	1140
Qy	1141	CCGCACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGCACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
Qy	1201	GCACAAATTCATGTTTGACAGCTTATCATCGCATGCAACGCTGCAACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATTCATGTTTGACAGCTTATCATCGCATGCAACGCTGCAACCAATGCTTCTGGCG	1260
Qy	1261	TCAGCAGCCATCGAAAGCTGTGTATGCTGTGCGAGCTGTAATAACATGCAATAATTCG	1320
Db	1261	TCAGCAGCCATCGAAAGCTGTGTATGCTGTGCGAGCTGTAATAACATGCAATAATTCG	1320
Qy	1321	TGTCGCTCAAGCGCACCTCCGGTCTGGATAATGTTTTTTTGGCCGACATCAACCGGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCGGTCTGGATAATGTTTTTTTGGCCGACATCAACCGGTT	1380
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTAATGTGTGGA	1440
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTAATGTGTGGA	1440
Qy	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAAACAGCCAGTCCGTTTGGTGGTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAAACAGCCAGTCCGTTTGGTGGTTCACGA	1500
Qy	1501	GCATTTACCAACAAAGGACCATAGATTATGAAACCTGGAAGGTAACCTGGTAATCTCG	1560
Db	1501	GCATTTACCAACAAAGGACCATAGATTATGAAACCTGGAAGGTAACCTGGTAATCTCG	1560
Qy	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTGGTAAAGGTAATTCGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTGGTAAAGGTAATTCGAAAGAT	1620
Qy	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGTAATCCCAAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGTAATCCCAAGGTT	1680
Qy	1681	GGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTGGTGCTAC	1740
Db	1681	GGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTGGTGCTAC	1740
Qy	1741	GCTCAATCTGGCGTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGCAACGCTGTAAT	1800
Db	1741	GCTCAATCTGGCGTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGCAACGCTGTAAT	1800
Qy	1801	CCGTTTACCTGGATGCCGATGCTTAAACGGCAAGCTGATTTGCTTACCAGTCCGCTGTT	1860
Db	1801	CCGTTTACCTGGATGCCGATGCTTAAACGGCAAGCTGATTTGCTTACCAGTCCGCTGTT	1860
Qy	1861	GAGCGTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCAAAAAACCTGGGAA	1920
Db	1861	GAGCGTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCAAAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTATCTTCACTTGGCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTATCTTCACTTGGCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAAAACGGCAAGTACGACATTAAGACGTTGAGCGGTGGATAACCGCTGGCGGAAACG	2100
Db	2041	TATGAAAAACGGCAAGTACGACATTAAGACGTTGAGCGGTGGATAACCGCTGGCGGAAACG	2100
Qy	2101	GGTCTGACCTTCTGTTGCTGACCTGATTAATAACAAACATGAAATGACAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGTTGCTGACCTGATTAATAACAAACATGAAATGACAGACACCGATTAC	2160

2161	Qy	TCCATCGCAGAGCTGCTTTTAAATAAAGGGCAAAACAGCGATGACCATCAACCGGCCCGTGG	2222
2161	Db		2222
		TCCATCGCAGAGCTGCTTTTAAATAAAGGGCAAAACAGCGATGACCATCAACCGGCCCGTGG	
2221	Qy	GCATGGTCCACATCGACACAGCAGAAAGTGAATTTATGGTGTAAAGGTACTGCCGACCTTC	2280
2221	Db	GCATGGTCCACATCGACACAGCAGAAAGTGAATTTATGGTGTAAAGGTACTGCCGACCTTC	2280
2281	Qy	AAGGGTCAACCATCCAAACCGTTGGTGGCGTGTGAGCGCAGGTATTTAAACGCGCCAGT	2340
2281	Db	AAGGGTCAACCATCCAAACCGTTGGTGGCGTGTGAGCGCAGGTATTTAAACGCGCCAGT	2340
2341	Qy	CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACCTATCTGTGACTGATGAAGGTCTG	2400
2341	Db	CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACCTATCTGTGACTGATGAAGGTCTG	2400
2401	Qy	GAAGCGGTTAATAAGACAAACCCCTGGGTGCCGTAGCGTGAAGTCTTACGAGGAAGAG	2460
2401	Db	GAAGCGGTTAATAAGACAAACCCCTGGGTGCCGTAGCGTGAAGTCTTACGAGGAAGAG	2460
2461	Qy	TTGCGGAAGAATCCACGCTATTGCCGCGCACCATGTGAAAACGCCCAAGAGGTGAATCATG	2520
2461	Db	TTGCGGAAGAATCCACGCTATTGCCGCGCACCATGTGAAAACGCCCAAGAGGTGAATCATG	2520
2521	Qy	CCGAACATCCGCGAGATGTCGCTTTTCTTGGTATGCCGTGGCTACTCGCGGTGATCAACGCC	2580
2521	Db	CCGAACATCCGCGAGATGTCGCTTTTCTTGGTATGCCGTGGCTACTCGCGGTGATCAACGCC	2580
2581	Qy	GCCAGCGGTGCTCAGACTGTTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
2581	Db	GCCAGCGGTGCTCAGACTGTTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
2641	Qy	AACAACAAACAATAACAATAACAACAACTCTGGGATCGAGGGAAGATTTTCAGAATTC	2700
2641	Db	AACAACAAACAATAACAATAACAACAACTCTGGGATCGAGGGAAGATTTTCAGAATTC	2700

RESULT 2

```

US-09-096-570-1
; Sequence 1, Application US/09096570
; Patent No. 6018020
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Huret, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/096,570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522

```

QY 841 TCGCGCCATTACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGTAGTGGGATACG 900
Db |||||
QY 841 TCGCGCCATTACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGTAGTGGGATACG 900
Db |||||
QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCGCCGTTAAACACCATCAAAACAGGATTTTC 960
Db |||||
QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCGCCGTTAAACACCATCAAAACAGGATTTTC 960
Db |||||
QY 961 GCCTGCTGGGCAACACAGCTGGACCGCTTCTGCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db |||||
QY 961 GCCTGCTGGGCAACACAGCTGGACCGCTTCTGCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db |||||
QY 1021 AGGGCAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAACACACCTGGCGCCCAATA 1080
Db |||||
QY 1021 AGGGCAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAACACACCTGGCGCCCAATA 1080
Db |||||
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGAGTGGAGTGGAGTGG 1140
Db |||||
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGAGTGGAGTGGAGTGG 1140
Db |||||
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGAGTGGAGTGGAGTGG 1200
Db |||||
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGAGTGGAGTGGAGTGG 1200
Db |||||
QY 1201 GCACAAATCTCATGTTTGCAGCTTATCATCGACTGCAACGCAATTAATGAGTGGAGTGGAGTGG 1260
Db |||||
QY 1201 GCACAAATCTCATGTTTGCAGCTTATCATCGACTGCAACGCAATTAATGAGTGGAGTGGAGTGG 1260
Db |||||
QY 1261 TCAGGACGCAATCGGAGCTGTGTATGGCTGTGCGAGTGGAGTGGAGTGGAGTGGAGTGG 1320
Db |||||
QY 1261 TCAGGACGCAATCGGAGCTGTGTATGGCTGTGCGAGTGGAGTGGAGTGGAGTGGAGTGG 1320
Db |||||
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTCGCGCGACATCATACAGCGTT 1380
Db |||||
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTCGCGCGACATCATACAGCGTT 1380
Db |||||
QY 1381 CTGGCAATATCTGGAATGAGCTGTGGAACAATTAATCATCGCTCGTATATGTTGGA 1440
Db |||||
QY 1381 CTGGCAATATCTGGAATGAGCTGTGGAACAATTAATCATCGCTCGTATATGTTGGA 1440
Db |||||
QY 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTTCGTTAGTGGTTCACGA 1500
Db |||||
QY 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTTCGTTAGTGGTTCACGA 1500
Db |||||
QY 1501 GCATTCACCAACAGGACCATAGATTATGAAACTGAAGAAGTAACTGGTAATCTGG 1560
Db |||||
QY 1501 GCATTCACCAACAGGACCATAGATTATGAAACTGAAGAAGTAACTGGTAATCTGG 1560
Db |||||
QY 1561 ATTAAACGGCGATAAAGGCTATACCGGTCTCGTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620
Db |||||
QY 1561 ATTAAACGGCGATAAAGGCTATACCGGTCTCGTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620
Db |||||
QY 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT 1680
Db |||||
QY 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT 1680
Db |||||
QY 1681 CGCGCAATCGCGATGCGCTGACATTAATCTCTGGGCAACAGCGCTTGGTGGCTAC 1740
Db |||||
QY 1681 CGCGCAATCGCGATGCGCTGACATTAATCTCTGGGCAACAGCGCTTGGTGGCTAC 1740
Db |||||
QY 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCGCGAACAAGGTTTCAGGACAAGCTGTAT 1800
Db |||||
QY 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCGCGAACAAGGTTTCAGGACAAGCTGTAT 1800
Db |||||
QY 1801 CCGTTTACCTGGGATGCGTAGCTTCAACGGCAAGCTGATGTTTACCGGATCGCTGTT 1860
Db |||||
QY 1801 CCGTTTACCTGGGATGCGTAGCTTCAACGGCAAGCTGATGTTTACCGGATCGCTGTT 1860
Db |||||
QY 1861 GAACGCTTATCGTGTATTAACAAAGATCTGCTGCCGAACCGCGCAAAAACCTGGAA 1920
Db |||||
QY 1861 GAACGCTTATCGTGTATTAACAAAGATCTGCTGCCGAACCGCGCAAAAACCTGGAA 1920
Db |||||

QY 1921 GAGATCCGGCGCTGGATTAAGACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Db |||||
QY 1921 GAGATCCGGCGCTGGATTAAGACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Db |||||
QY 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATGCTGCTGACGGGGGTATGCGTTCAAG 2040
Db |||||
QY 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATGCTGCTGACGGGGGTATGCGTTCAAG 2040
Db |||||
QY 2041 TATGAAAACCGCAAGTACGACATTAAGAGACGTGGCGGTGATACGCTGGCGGAAAGCG 2100
Db |||||
QY 2041 TATGAAAACCGCAAGTACGACATTAAGAGACGTGGCGGTGATACGCTGGCGGAAAGCG 2100
Db |||||
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACACATGATGACAGACCGGATTAC 2160
Db |||||
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACACATGATGACAGACCGGATTAC 2160
Db |||||
QY 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACACGCGATGACCATCAACGGCGCGGTG 2220
Db |||||
QY 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACACGCGATGACCATCAACGGCGCGGTG 2220
Db |||||
QY 2221 GCATGTTCCAAACATCGACACAGCAAAAGTAAATATGTTGTAAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2221 GCATGTTCCAAACATCGACACAGCAAAAGTAAATATGTTGTAAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2281 AAGGTCACCAATCCAAACCGTTTCTGTTGGCGTCTGAGCGCAGGTATTAACGGCGCGCAGT 2340
Db |||||
QY 2281 AAGGTCACCAATCCAAACCGTTTCTGTTGGCGTCTGAGCGCAGGTATTAACGGCGCGCAGT 2340
Db |||||
QY 2341 CCGAAACAAAGAGTGGCAAAAGATTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db |||||
QY 2341 CCGAAACAAAGAGTGGCAAAAGATTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db |||||
QY 2401 GAAGCGGTTTAAATAAGCAACACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db |||||
QY 2401 GAAGCGGTTTAAATAAGCAACACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db |||||
QY 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATCGAAAAACCGCCAGAAAGGTGAAATCATG 2520
Db |||||
QY 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATCGAAAAACCGCCAGAAAGGTGAAATCATG 2520
Db |||||
QY 2521 CCGAAACATCCGCGAGATGTCGCTTTCTGTTGATGAGCGTGGTGGTGGTGGTGGTGGTGG 2580
Db |||||
QY 2521 CCGAAACATCCGCGAGATGTCGCTTTCTGTTGATGAGCGTGGTGGTGGTGGTGGTGGTGG 2580
Db |||||
QY 2581 GCAGCGGCTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCAGACTAAATTCGAGCTCG 2640
Db |||||
QY 2581 GCAGCGGCTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCAGACTAAATTCGAGCTCG 2640
Db |||||
QY 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAAAGGATTTTCAGAAATTC 2700
Db |||||
QY 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAAAGGATTTTCAGAAATTC 2700
Db |||||

RESULT 3

US-09-265-617B-1
; Sequence 1, Application US/09265617B
; Patent No. 6372883
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R.
; APPLICANT: Hurst, David N.
; APPLICANT: Jones, Philip S.
; APPLICANT: Kay, Paul B.
; APPLICANT: Raynham, Tony M.
; APPLICANT: Wilson, Francis X.
; TITLE OF INVENTION: Antiviral Medicaments
; FILE REFERENCE: 20052 antiviral medicaments
; CURRENT APPLICATION NUMBER: US/09/265,617B
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: GB9806815.8
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1

; LENGTH: 7475									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: Vector and									
; OTHER INFORMATION: Gene Fragments									
US-09-265-617B-1									
Query Match 81.8%; Score 2700; DB 3; Length 7475;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CCGACACCATCGAATGCTGCAAAACCTTTCCGGGTATGCGCATGATAGCGCCGGAAGAGA	60						
Db	1	CCGACCATCATGAAATGGTGCAAAACCTTTCCGGGTATGCGCATGATAGCGCCGGAAGAGA	60						
Qy	61	GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG	120						
Db	61	GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG	120						
Qy	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180						
Db	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180						
Qy	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATTTACATTTCCCAACCGCTGGCAC	240						
Db	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATTTACATTTCCCAACCGCTGGCAC	240						
Qy	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTTGCCACCTCCAGTCTGGCCCTGC	300						
Db	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTTGCCACCTCCAGTCTGGCCCTGC	300						
Qy	301	ACGCGCGTGTGCAATTTGTGCGGCGATTAATCTTCGCGCGGATCAACTGGGTGCCAGG	360						
Db	301	ACGCGCGTGTGCAATTTGTGCGGCGATTAATCTTCGCGCGGATCAACTGGGTGCCAGG	360						
Qy	361	TGGTGTTCGATGTGAACGAAAGCGGCTGAAAGCTGTAAAGCGGCGGTGCACAATC	420						
Db	361	TGGTGTTCGATGTGAACGAAAGCGGCTGAAAGCTGTAAAGCGGCGGTGCACAATC	420						
Qy	421	TTCTCGCGCAACCGCTCAGTGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCA	480						
Db	421	TTCTCGCGCAACCGCTCAGTGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCA	480						
Qy	481	TTGCTGTGGAAGCTCCCTGCACTAATTTTCCGGCGTATTTCTTGATGCTCTGACCCAGA	540						
Db	481	TTGCTGTGGAAGCTCCCTGCACTAATTTTCCGGCGTATTTCTTGATGCTCTGACCCAGA	540						
Qy	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
Db	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
Qy	601	TGGTCGCATTTGGTCCACAGCAAAATCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG	660						
Db	601	TGGTCGCATTTGGTCCACAGCAAAATCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG	660						
Qy	661	CGCGTCTGCGTCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGCGTCTGCGTCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Qy	721	CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780						
Db	721	CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780						
Qy	781	ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGCAA	840						
Db	781	ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGCAA	840						
Qy	841	TGCGGCGCATTAACCGAGTCCGGGCTGCGGTTTGGTGGGATATCTCGGTAGTGGGATAG	900						
Db	841	TGCGGCGCATTAACCGAGTCCGGGCTGCGGTTTGGTGGGATATCTCGGTAGTGGGATAG	900						
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAAGGATTTTC	960						

Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAAGGATTTTC	960						
Qy	961	GCCTGCTGGGCAAAACAGCGTGAACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA	1020						
Db	961	GCCTGCTGGGCAAAACAGCGTGAACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA	1020						
Qy	1021	AGGCAATCAGCTGTGCCGCTCTCACTGCTGAAGAAAGAAACCAACCTGGCGCCAATA	1080						
Db	1021	AGGCAATCAGCTGTGCCGCTCTCACTGCTGAAGAAAGAAACCAACCTGGCGCCAATA	1080						
Qy	1081	CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGGCAGACAGGTTT	1140						
Db	1081	CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGGCAGACAGGTTT	1140						
Qy	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200						
Db	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200						
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGG	1260						
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGG	1260						
Qy	1261	TCAGCAGCCATCGGAAGCTGTGCTGTGAGCTGTCAGGTTCGTAATCACTGCATAATTTCG	1320						
Db	1261	TCAGCAGCCATCGGAAGCTGTGCTGTGAGCTGTCAGGTTCGTAATCACTGCATAATTTCG	1320						
Qy	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATACGGTT	1380						
Db	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATACGGTT	1380						
Qy	1381	CTGGCAATATCTGAAATGAGCTGTTCACATTAATCATCGGCTCGTATATGTTGGA	1440						
Db	1381	CTGGCAATATCTGAAATGAGCTGTTCACATTAATCATCGGCTCGTATATGTTGGA	1440						
Qy	1441	ATTGTGAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500						
Db	1441	ATTGTGAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500						
Qy	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAGGTAAATCTGGTAATCTGG	1560						
Db	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAGGTAAATCTGGTAATCTGG	1560						
Qy	1561	ATTTAAGCGGATAAAGGCTATAACGCTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620						
Db	1561	ATTTAAGCGGATAAAGGCTATAACGCTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620						
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680						
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680						
Qy	1681	CGCGCAACTGGCGATGGCCCTGCACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Db	1681	CGCGCAACTGGCGATGGCCCTGCACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Qy	1741	GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGACAAGCTGTAT	1800						
Db	1741	GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGACAAGCTGTAT	1800						
Qy	1801	CCGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCCGCTTT	1860						
Db	1801	CCGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCCGCTTT	1860						
Qy	1861	GAAAGCTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Db	1861	GAAAGCTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980						
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980						
Qy	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGCTGACGGGGTTATCGGTTCAAG	2040						

1981	CTGCAAGAACGGTACTTCAACCTGGCCGCTGATTGCTCTGCTGACGGGGGTATTGCGTTCAAG	2040
2041	TATGAAAAACGGCAAGTAGTCAGCATTTAAAGACGTTGGCGCTGGATACGCTGGCGCGAAAGCG	2100
2041	TATGAAAAACGGCAAGTAGTCAGCATTTAAAGACGTTGGCGCTGGATTAACGCTGGCGCGAAAGCG	2100
2101	GGTCTGACCTTCTGGTTGACCTGATTTAAAAAACAACACATGAATGAGACACCGGATTAC	2160
2101	GGTCTGACCTTCTGGTTGACCTGATTTAAAAAACAACACATGAATGAGACACCGGATTAC	2160
2161	TCCATCGCAGAAAGCTGCCTTTTATAAAGGCGAAACACGCGATGACCATCAACGCGCCGTGG	2220
2161	TCCATCGCAGAAAGCTGCCTTTTATAAAGGCGAAACACGCGATGACCATCAACGCGCCGTGG	2220
2221	GCATGTTCCAACTCGACACACAGCAAAAGTGAATTTATGGTTAAACGGTACTGCCGACCTTC	2280
2221	GCATGTTCCAACTCGACACACAGCAAAAGTGAATTTATGGTTAAACGGTACTGCCGACCTTC	2280
2281	AAGGGTCAACCATCTCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTTAACGCGCGCAGT	2340
2281	AAGGGTCAACCATCTCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTTAACGCGCGCAGT	2340
2341	CCGAAACAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGTGACTGATGAAGGTCGTG	2400
2341	CCGAAACAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGTGACTGATGAAGGTCGTG	2400
2401	GAAGCGTTTAATAAGACAACACCGCTGGTGCCGTAGCGCTGAAGTCTTACGAGGAGAG	2460
2401	GAAGCGTTTAATAAGACAACACCGCTGGTGCCGTAGCGCTGAAGTCTTACGAGGAGAG	2460
2461	TTGGCGAAAGATCCACGTATTTCGCCGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG	2520
2461	TTGGCGAAAGATCCACGTATTTCGCCGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG	2520
2521	CCGAAACATCCGCGAGATGTCGCTTTCTGTTATGCGTGGGTAATCTGGGTGATCAACGCC	2580
2521	CCGAAACATCCGCGAGATGTCGCTTTCTGTTATGCGTGGGTAATCTGGGTGATCAACGCC	2580
2581	GCCAGCGGTGCTCAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
2581	GCCAGCGGTGCTCAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
2641	AACAACAACAATTAACAATAACAACAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC	2700
2641	AACAACAACAATTAACAATAACAACAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC	2700

RESULT 4

US-09-027-169-3/C
Sequence 3, Application US/09027169
Patent No. 6420524
GENERAL INFORMATION:
APPLICANT: CRAIG, NANCY L
TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
TITLE OF INVENTION: ATP-DEPENDENT TRANSCRIPTION PROTEINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Anne Brown (Alston & Bird, LLP)
STREET: 3605 Glenwood Ave.
CITY: Raleigh
STATE: NC
COUNTRY: USA
ZIP: 27608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/027,169
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

|||||
784 GAGGCATCGTTCACATGCGATGCTGGTGGCAACGATCAGATGGCGCTGGGCGCAATG 725
QY CGCGCATTTACCGAGTCGCGGCTGGCGCTGGTGGCGGATATCTCGGTAGTGGGATACGAC 902
Db CGCGCATTTACCGAGTCGCGGCTGGCGCTGGTGGCGGATATCTCGGTAGTGGGATACGAC 665
QY GATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCAATCAACAGAGATTTTCG 962
Db GATACCGAAGACAGCTCATGTTATATCCGCGGTTCAACCAATCAACAGAGATTTTCG 605
QY CTGCTGGGCGAAACAGAGCTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db CTGCTGGGCGAAACAGAGCTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGAAG 545
QY GGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAGAAACACACCTTGGCGCCCAATACG 1082
Db GGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAGAAACACACCTTGGCGCCCAATACG 485
QY CAACCGGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAGCTGGACGACAGGTTTCC 1142
Db CAACCGGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAGCTGGACGACAGGTTTCC 425
QY CGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAGGC 1202
Db CGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGC-----GC 376
QY ACAATTCTCATGTTTGACAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCGTC 1262
Db GAAATTGATCTGGTTTGACAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCGTC 316
QY AGGAGCCATCGGAAGCTGTGGTATGGTGTGCGAGTGGTAAATCACTGCAATAATTCGTG 1322
Db AGGAGCCATCGGAAGCTGTGGTATGGTGTGCGAGTGGTAAATCACTGCAATAATTCGTG 256
QY TCGCTCAAGGCGCACTCCGCTTCTGGATAATGTTTTTGGCGCGACATCAACGGTTCT 1382
Db TCGCTCAAGGCGCACTCCGCTTCTGGATAATGTTTTTGGCGCGACATCAACGGTTCT 196
QY GGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCAT- CGGCTCGTATAATGTGGAA 1441
Db GGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCATCAATCAATCATCGGCTCGTATAATGTGGAA 136
QY TTGTGAGCGGATAACAATTTACACAGGAAACAGCC 1477
Db TTGTGAGCGGATAACAATTTACACAGGAAACAGAC 100

RESULT 5
US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE2
US-09-640-882-2

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
QY 3 GACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCCGAAAGAGT 62
Db 866 GACACCATCGAATGGCGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCCGAAAGAGT 925
QY 63 CAATTCAGGGTGGTGAATGTGAAACCCAGTAAAGTTATACGATGTGCGAGATGATGCCGGT 122
Db 926 CAATTCAGGGTGGTGAATGTGAAACCCAGTAAAGTTATACGATGTGCGAGATGATGCCGGT 985
QY 123 GTCTCTTATCAGACGGTTTCCGCGTGGTGAACAGGCGCCAGCTTTCTCGGAAACG 182
Db 986 GTCTCTTATCAGACGGTTTCCGCGTGGTGAACAGGCGCCAGCTTTCTCGGAAACG 1045
QY 183 CGGGAAGAGTGAAGCGGCGATGGCGGAGCTGAATTAATCTCGCGCGATCAATCGGTCGCAAA 242
Db 1046 CGGGAAGAGTGAAGCGGCGATGGCGGAGCTGAATTAATCTCGCGCGATCAATCGGTCGCAAA 1105
QY 243 CAATTCAGGGTGAAGTGAATGTGAAACCCAGTAAAGTTATACGATGTGCGAGTCTGCCCTGAC 302
Db 1106 CAATTCAGGGTGAAGTGAATGTGAAACCCAGTAAAGTTATACGATGTGCGAGTCTGCCCTGAC 1165
QY 303 CGCGCTGCGCAAAATGTTCGCGCGATTAATTAATCTCGCGCGATCAATCGGTCGCGAGCGTG 362
Db 1166 CGCGCTGCGCAAAATGTTCGCGCGATTAATTAATCTCGCGCGATCAATCGGTCGCGAGCGTG 1225
QY 363 GTGGTGTGATGATGAGAACGCGGCTGCAAGCTGTAAGCGGCGGTGCACAACTTT 422
Db 1226 GTGGTGTGATGATGAGAACGCGGCTGCAAGCTGTAAGCGGCGGTGCACAACTTT 1285
QY 423 CTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCAATCTCGCTGGATGACAGGATGCCATT 482
Db 1286 CTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCAATCTCGCTGGATGACAGGATGCCATT 1345
QY 483 GCTGTGAGAGCTGCTGCACATAATGTTCCGCGGTATTTCTTGATGTCTTGACAGACA 542
Db 1346 GCTGTGAGAGCTGCTGCACATAATGTTCCGCGGTATTTCTTGATGTCTTGACAGACA 1405
QY 543 CCCATCAACAGTATTTCTCCATGAAGACGTCAGCGACTGGGCGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTTCTCCATGAAGACGTCAGCGACTGGGCGTGGAGCATCTG 1465
QY 603 CTCGATTGGGTCAACAGCAAAATCGCGTGTGTAGCGGCGCCATTAAGTTCTGTCTCGGCG 662
Db 1466 CTCGATTGGGTCAACAGCAAAATCGCGTGTGTAGCGGCGCCATTAAGTTCTGTCTCGGCG 1525
QY 663 CGTCTGCGTCTGGCTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 722
Db 1526 CGTCTGCGTCTGGCTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 1585
QY 723 GAACGGGAAGCGACTGGAGTGGCATGTCGCGTTTCAACAAACCATCAAAATGCTGAAT 782
Db 1586 GAACGGGAAGCGACTGGAGTGGCATGTCGCGTTTCAACAAACCATCAAAATGCTGAAT 1645
QY 783 GAGGCGATCGTTTCCACATCGCGATGCTGTTGTCACCAACGATCAGATGGCGCTGGGCGCAATG 842
Db 1646 GAGGCGATCGTTTCCACATCGCGATGCTGTTGTCACCAACGATCAGATGGCGCTGGGCGCAATG 1705
QY 843 CGCGCATTAACGAGTCCGCGCTGCGGCTGGTGGTGGATATCTCGGTAGTGGATACGAC 902
Db 1706 CGCGCATTAACGAGTCCGCGCTGCGGCTGGTGGTGGATATCTCGGTAGTGGATACGAC 1765
QY 903 GATACCGAAGACGCTCATGTTATATCCGCGGTTAACCAACCATCAAAACAGATTTTTCG 962
Db 1766 GATACCGAAGACGCTCATGTTATATCCGCGGTTAACCAACCATCAAAACAGATTTTTCG 1825
QY 963 CTGCTGGGCGAAACAGGCTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGAAG 1022
Db 1826 CTGCTGGGCGAAACAGGCTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGAAG 1885
QY 1023 GGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAGAAACCAACCTCGCGCCCAATACG 1082

```
Db 1886 GSCAATCAGCTGTGCGCGTCTCACTGGTGAAAGAAAACCACTCGCGCCCAATAGC 1945
Qy 1083 CAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCGAGTGGCGAGCAGGTTTCC 1142
Db 1946 CAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCGAGTGGCGAGCAGGTTTCC 2005
Qy 1143 CCACTGGAAGCGGCGAGTGGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTAGGC 1202
Db 2006 CCACTGGAAGCGGCGAGTGGAGCGCAACCAATTAATGTAGTTAG--CGGAATTGATC 2063
Qy 1203 ACAATTTCTCATGTTGACAGCTTATCATCGACTGACGGTGCAACCAATGCTTCTGGCGTC 1262
Db 2064 TGAATTTCTCATGTTGACAGCTTATCATCGACTGACGGTGCAACCAATGCTTCTGGCGTC 2123
Qy 1263 AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGGTAATCACTGCATTAATTCGTG 1322
Db 2124 AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGGTAATCACTGCATTAATTCGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCCGTCTGGATTAATTTTTTTTGGCGCGACATCAATACGGTTCT 1382
Db 2184 TCGCTCAAGGCGCACTCCCGTCTGGATTAATTTTTTTTGGCGCGACATCAATACGGTTCT 2243
Qy 1383 GGCATAATTTCT 1394
Db 2244 GGCATAATTTCT 2255

RESULT 6
US-09-640-882-3
; Sequence 3, Application US/09640882
; Patent No. 672042
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE3
US-09-640-882-3

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 3 GACACCATCGAATGGTGCAGAAACCTTTTCGCGTATGGCATGATAGCGCCCGGAAGAGAGT 62
Db 866 GACACCATCGAATGGCGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGT 925
Qy 63 CAATTTCAGGTGGTGAATGTGAACCAAGTAAACGTTATACGATGTCGAGAGTATGCCGTT 122
Db 926 CAATTTCAGGTGGTGAATGTGAACCAAGTAAACGTTATACGATGTCGAGAGTATGCCGTT 985
Qy 123 GTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCCAGCCACGTTTCTCGGAAACG 182
Db 986 GTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCCAGCCACGTTTCTCGGAAACG 1045
Qy 183 CGGGAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCCAAACCGGTGGCAAA 242
Db 1046 CGGGAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCCAAACCGGTGGCAAA 1105
```

```
Qy 243 CAACCTGGCGGCAACAGTCTGTTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGCAC 302
Db 1106 CAACCTGGCGGCAACAGTCTGTTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGCAC 1165
Qy 303 GCGCGTTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCCGATCAATCGGGTGCAGCGTG 362
Db 1166 GCGCGTTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCCGATCAATCGGGTGCAGCGTG 1225
Qy 363 GTGGTTCGATGTGAGAAACGAAGCGGCTGGAAGCTGTAAAGCGGCGGTGCACAAATCTT 422
Db 1226 GTGGTTCGATGTGAGAAACGAAGCGGCTGGAAGCTGTAAAGCGGCGGTGCACAAATCTT 1285
Qy 423 CTGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACAGGATGCCATT 482
Db 1286 CTGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACAGGATGCCATT 1345
Qy 483 GCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGTATGTTCTCTGACAGACA 542
Db 1346 GCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGTATGTTCTCTGACAGACA 1405
Qy 543 CCCATCAACAGTATTTTCTCCATGAAGACGTTACGGCACTGGCGGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTTTCTCCATGAAGACGTTACGGCACTGGCGGTGGAGCATCTG 1465
Qy 603 GTGCGATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAAGTCTGTCGCGG 662
Db 1466 GTGCGATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAAGTCTGTCGCGG 1525
Qy 663 CGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCGCATAGCG 722
Db 1526 CGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCGCATAGCG 1585
Qy 723 GAACGGGAAGCGCACTGGAGTGCATGTCGGTTTCAACAAACCAATGCAATGCTGAAT 782
Db 1586 GAACGGGAAGCGCACTGGAGTGCATGTCGGTTTCAACAAACCAATGCAATGCTGAAT 1645
Qy 783 GAGGCGATCGTTCCTCCACTGCGATGCTGTTGCGCAACGATCAGATGCGCTGGCGCAATG 842
Db 1646 GAGGCGATCGTTCCTCCACTGCGATGCTGTTGCGCAACGATCAGATGCGCTGGCGCAATG 1705
Qy 843 CGCGCATTAACGAGTCCGGGCTGGCGGTGTTGGTGCGGATATCTCGGTAGTGGGATACGAC 902
Db 1706 CGCGCATTAACGAGTCCGGGCTGGCGGTGTTGGTGCGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTCG 1825
Qy 963 CTGCTGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGAAG 1022
Db 1826 CTGCTGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGAAG 1885
Qy 1023 GGAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAAACCAACCTGGCGCCCAATAGC 1082
Db 1886 GGAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAAACCAACCTGGCGCCCAATAGC 1945
Qy 1083 CAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATCAGCTGGCAGCAGAGTTTCC 1142
Db 1946 CAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATCAGCTGGCAGCAGAGTTTCC 2005
Qy 1143 CCACTGGAAGCGGCGAGTGGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTAGGC 1202
Db 2006 CCACTGGAAGCGGCGAGTGGAGCGCAACCAATTAATGTAGTTAG--CGCAATTTGATC 2063
Qy 1203 ACAATTTCTCATGTTGACAGCTTATCATCGACTGACGGTGCAACCAATGCTTCTGGCGTC 1262
Db 2064 TGAATTTCTCATGTTGACAGCTTATCATCGACTGACGGTGCAACCAATGCTTCTGGCGTC 2123
Qy 1263 AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGGTAATCACTGCATTAATTCGTG 1322
Db 2124 AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGGTAATCACTGCATTAATTCGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCCGTCTGGATTAATTTTTTTTGGCGCGACATCAATACGGTTCT 1382
```

Db 2184 TCGCTCAAGGCGCATCCCGTTCTGGATAATGTTTTTGGCCGCAATCATACCGTTCT 2243
QY 1383 GGCAAAATATCT 1394
Db 2244 GGCAAAATATCT 2255

RESULT 7

US-08-778-717-5
; Sequence 5, Application US/08778717
; Patent No. 6602689
; GENERAL INFORMATION:
; APPLICANT: UENO, EIICHI
; APPLICANT: NOBUYUKI, FUJII
; APPLICANT: OKADA, MASAHISA
; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN
; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR
; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/778,717
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 352225/1995
; FILING DATE: 28-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2084-031-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4557 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; ORIGINAL SOURCE:
; ORGANISM: E. COLI
; STRAIN: BL21 (DE3)
; PUBLICATION INFORMATION:
; AUTHORS: NOBUYUKI FUJII ET AL,
; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED
; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF
; TITLE: EXPRESSING SAID FUSED PROTEIN
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557

US-08-778-717-5
Query Match 36.5%; Score 1204; DB 4; Length 4557;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACCATCAATGGTGCAAAACCTTTCGGCGGTATGGCATGATAGCCCGGAAGAGA 60
Db 2563 CCGACACCATCAATGGTGCAAAACCTTTCGGCGGTATGGCATGATAGCCCGGAAGAGA 2622

QY 61 GTCAATTCAAGGTGGTGAATGTAAGCAAGTAAAGTTATACGATGTCGCAGAGTATCCG 120
Db 2623 GTCAATTCAAGGTGGTGAATGTAAGCAAGTAAAGTTATACGATGTCGCAGAGTATCCG 2682
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCAGCCAGTCTTCTCGCAAAA 180
Db 2683 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCAGCCAGTCTTCTCGCAAAA 2742
QY 181 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 2743 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 2802
QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 2803 AACAACTGGCGGCAAAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGCCCTGC 2862
QY 301 ACGGCGCTCGCAAAATTCGCGCGCATTAATCTCGGCGCGATCAATCTGGGTGCCAGCG 360
Db 2863 ACGGCGCTCGCAAAATTCGCGCGCATTAATCTCGGCGCGATCAATCTGGGTGCCAGCG 2922
QY 361 TGGTGGTGTGATGGTAGAAGCGCGCTGGAAGCCTGTAAAGCGCGGTGCAAAATC 420
Db 2923 TGGTGGTGTGATGGTAGAAGCGCGCTGGAAGCCTGTAAAGCGCGGTGCAAAATC 2982
QY 421 TTCTCGGCAACCGCTCAGTGGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
Db 2983 TTCTCGGCAACCGCTCAGTGGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3042
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db 3043 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 3102
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 3103 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 3162
QY 601 TGGTGGCATTTGGGTACCAAGCAAAATCGCGCTGTTAGCGGCGCCATTAATGTTCTGTCTGG 660
Db 3163 TGGTGGCATTTGGGTACCAAGCAAAATCGCGCTGTTAGCGGCGCCATTAATGTTCTGTCTGG 3222
QY 661 CGCGTCTCGCTCTGGCTGGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
Db 3223 CGCGTCTCGCTCTGGCTGGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3282
QY 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 3283 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 3342
QY 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGCCAACGATCAGATGGCGTGGCGGCA 840
Db 3343 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGCCAACGATCAGATGGCGTGGCGGCA 3402
QY 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAGG 900
Db 3403 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAGG 3462
QY 901 ACGATACCGGAAGACAGCTCATGTTATATCCGCGGTTTAAACCAACCATCAACAGGATTTTC 960
Db 3463 ACGATACCGGAAGACAGCTCATGTTATATCCGCGGTTTAAACCAACCATCAACAGGATTTTC 3522
QY 961 GCCTGCTGGGGCAAAACAGCGTGGAGCCGTTGTCGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3523 GCCTGCTGGGGCAAAACAGCGTGGAGCCGTTGTCGCAACTCTCTCAGGGCCAGGCGGTGA 3582
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCAGCTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 3583 AGGGCAATCAGCTGTTGCCCGTCTCAGCTGGTGAAGAAACCAACCTGGCGCCCAATA 3642
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCAGAGTTT 1140
Db 3643 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCAGAGTTT 3702
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTGTAGTCTACTCATTTAG 1200

Db 3703 CCCGACTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTACTCACTCATTTAG 3762
Qy 1201 GCAC 1204
Db 3763 GCAC 3766

RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerita, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148.675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Isal, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3832 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-148-675A-2

Query Match 36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGGCAAAACCTTCCCGGTATGGCATGATAGCGCCCGGAAAGAGA 60
Db 3 CCGACACCATCGAATGGCGCAAAACCTTCCCGGTATGGCATGATAGCGCCCGGAAAGAGA 62

Qy 61 GTCAATTTCAGGGTGGTGAATGTGAACACAGTAAAGTATACGATGTCGACAGATATGCCG 120
Db 63 GTCAATTTCAGGGTGGTGAATGTGAACACAGTAAAGTATACGATGTCGACAGATATGCCG 122

Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGTAAAGTATACGATGTCGACAGATATGCCG 180
Db 123 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGTAAAGTATACGATGTCGACAGATATGCCG 182

Qy 181 CGCGGGAAGAGTGGAGCGCGGATGGCGAGCTGAATTATCATCCCAACGGGTGGCAC 240
Db 183 CGCGGGAAGAGTGGAGCGCGGATGGCGAGCTGAATTATCATCCCAACGGGTGGCAC 242

Qy 241 AACAACTGGCGGGCAACAGTCGTTGTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300

Db 243 AACAACTGGCGGGCAACAGTCGTTGCTGATTGGCGTTGGCACTCCAGTCTGGCCCTGC 302
Qy 301 ACGGCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAATCTGGGTGCCAGCG 360
Db 303 ACGGCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAATCTGGGTGCCAGCG 362

Qy 361 TGGTGGTGTTCGATGTTAGAACGAAAGCGGCTGCGAAGCTCTGTAAGCGCGGTGCACAATC 420
Db 363 TGGTGGTGTTCGATGTTAGAACGAAAGCGGCTGCGAAGCTCTGTAAGCGCGGTGCACAATC 422

Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACACGAGATGCCA 480
Db 423 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACACGAGATGCCA 482

Qy 481 TTGCTGTGGAGCTGCGCTGCACATAATGTTCCGGGTTATTTCTTGATGTTCTTGACGAGA 540
Db 483 TTGCTGTGGAGCTGCGCTGCACATAATGTTCCGGGTTATTTCTTGATGTTCTTGACGAGA 542

Qy 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTCGGCGTGGAGCATC 600
Db 543 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTCGGCGTGGAGCATC 602

Qy 601 TGGTCCGATTGGTTCACGACAAATCCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db 603 TGGTCCGATTGGTTCACGACAAATCCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 662

Qy 661 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATG 720
Db 663 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATG 722

Qy 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTCCTCAACAAACCATGCAATGCTGA 780
Db 723 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTCCTCAACAAACCATGCAATGCTGA 782

Qy 781 ATGAGGCATCGTTCCCATCGCATGTCGGATGTCGCAACAGATCAGATGGCGCTGGGCGCAA 840
Db 783 ATGAGGCATCGTTCCCATCGCATGTCGGATGTCGCAACAGATCAGATGGCGCTGGGCGCAA 842

Qy 841 TGGCGGCATTACCGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 843 TGGCGGCATTACCGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 902

Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC 960
Db 903 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC 962

Qy 961 GCCTGCTGGGGCAACACGAGGTGACCGCTTGCTGCAACTCTCTCAGGGCGCAGCGGTGA 1020
Db 963 GCCTGCTGGGGCAACACGAGGTGACCGCTTGCTGCAACTCTCTCAGGGCGCAGCGGTGA 1022

Qy 1021 AGGGCAATCAGCTGTTGGCGCTCTCACTGGTGAAGAAACACCCCTGGCGCCCAATA 1080
Db 1023 AGGGCAATCAGCTGTTGGCGCTCTCACTGGTGAAGAAACACCCCTGGCGCCCAATA 1082

Qy 1081 CGAAACCGCCTCTCCCGCGGTGGCGGATTCATTAAATGCAAGCTGGCACGACGAGTTT 1140
Db 1083 CGAAACCGCCTCTCCCGCGGTGGCGGATTCATTAAATGCAAGCTGGCACGACGAGTTT 1142

Qy 1141 CCGGACTGGAAAGCGGCGAGTGACCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1143 CCGGACTGGAAAGCGGCGAGTGACCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1202

Qy 1201 GCAC 1204
Db 1203 GCAC 1206

RESULT 9
US-08-487-283A-18/c
; Sequence 18, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.

APPLICANT: Matis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Squinto, Stephen P.
APPLICANT: Thomas, Thomas C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: 25 Science Park (Alexion)
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Cetrus 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)776-1790
TELEFAX: (203)772-3655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: PET Trc SOS/NI
DESCRIPTION: prokaryotic expression vector
US-08-487-283A-18

Query Match 36.3%; Score 1199.2; DB 3; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAACTTTCCGGGTATGCGATAGCGCCGGAAGAGA 60
DB 4810 CGGACACCATCGAATGGCGCAAAACCTTTCGGGTATGCGATAGCGCCGGAAGAGA 4751
QY 61 GTCATTCAGGTGGTGAATGTGAACCAAGTAAACCGTTATACGATCGCAGAGTATGCG 120
DB 4750 GTCATTCAGGTGGTGAATGTGAACCAAGTAAACCGTTATACGATCGCAGAGTATGCG 4691
QY 121 GTGTCTTTATCAGACCGTTTCCCGGTGGTGAACCAAGCGCCAGCCACGTTTCTCGAAAA 180
DB 4690 GTGTCTTTATCAGACCGTTTCCCGGTGGTGAACCAAGCGCCAGCCACGTTTCTCGAAAA 4631
QY 181 CGCGGAAAGTGAAGCGGATGGCGGAGCTGAATTACATTCGCAACCGCTGGGCAC 240
DB 4630 CGCGGAAAGTGAAGCGGATGGCGGAGCTGAATTACATTCGCAACCGCTGGGCAC 4571
QY 241 AACAACTCGCGGCAAAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 4570 AACAACTCGCGGCAAAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 4511
QY 301 ACGGCGCTCGCAAAATTTGTTCGGCGCATTAATAATCTCGCGCGCATCAACTGGGTGCCGCG 360

RESULT 10

PCT-US96-05611A-21/c
; Sequence 21, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Leonardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott

Db 4510 ACGCGCGTTCGCAAAATTTGTTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCGACGG 4451
QY 361 TGGTGGTGTTCGATGTAAGACGAAAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 4450 TGGTGGTGTTCGATGTAAGACGAAAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCAACAATC 4391
QY 421 TTCTCGCGCAACCGGTTCAGTGGGCTGATCAATACTATCCGCTGGATGACAGGATGCCA 480
Db 4390 TTCTCGCGCAACCGGTTCAGTGGGCTGATCAATACTATCCGCTGGATGACAGGATGCCA 4331
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db 4330 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
Db 4270 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 4211
QY 601 TGGTGGCATTTGGGTACACAGCAAAATCGCGCTGTGTAGCGGGCCCATTAAGTTCTGTCTCG 660
Db 4210 TGGTGGCATTTGGGTACACAGCAAAATCGCGCTGTGTAGCGGGCCCATTAAGTTCTGTCTCG 4151
QY 661 CGGCTCTGCGTCTGCGTGGCTGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4150 CGGCTCTGCGTCTGCGTGGCTGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
QY 721 CGGAAACGGGAAGCGGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 4090 CGGAAACGGGAAGCGGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCATCGCATGTGTTGGCAACGATCAGATGCGCTGGCGCGCA 840
Db 4030 ATGAGGGCATCGTTCCCATCGCATGTGTTGGCAACGATCAGATGCGCTGGCGCGCA 3971
QY 841 TGGCGGCAATACCGAGTCCGGCTGGCGTGGTGGCGATATCTCGTAGTGGGATAG 900
Db 3970 TGGCGGCAATACCGAGTCCGGCTGGCGTGGTGGCGATATCTCGTAGTGGGATAG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCATCAAAAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCATCAAAAGGATTTTC 3851
QY 961 GCCTCTGGGGCAAAACAGCGTGAACCGCTTGTGTCAACTCTCTCAGGCGCAGGCGTGA 1020
Db 3850 GCCTCTGGGGCAAAACAGCGTGAACCGCTTGTGTCAACTCTCTCAGGCGCAGGCGTGA 3791
QY 1021 AGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3790 AGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 3731
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAAATGAGCTGCGACGACAGGTTT 1140
Db 3730 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAAATGAGCTGCGACGACAGGTTT 3671
QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 3670 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607

APPLICANT: Nye, Steven H.
APPLICANT: Pelfrey, Clara M.
APPLICANT: Squinto, Stephen P.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: Modified Myelin Protein Molecules
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
COMPUTER: Macintosh Centris 610
OPERATING SYSTEM: System 7
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05611A
FILING DATE: 02-MAY-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,644
FILING DATE: May 2, 1995
APPLICATION NUMBER: 08/431,648
FILING DATE: May 2, 1995
APPLICATION NUMBER: 08/482,114
FILING DATE: June 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-129
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: PET Trc S05/NI
DESCRIPTION: prokaryotic expression vector
PCT-US96-05611A-21

Query Match 36.3%; Score 1199.2; DB 5; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGAACAACTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 4810 CGGACACCATCGAATGGCGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 4751
Qy 61 GTCAATTCAGGTCGTGATGTAAGCAACAGTAACTTATACGATTATACGATGTCGAGATGCGG 120
Db 4750 GTCAATTCAGGTCGTGATGTAAGCAACAGTAACTTATACGATGTCGAGATGTCGCG 4691
Qy 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACACGAGCGAGCTGAATTATCCCAACCGGTGGCAC 240
Db 4690 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACACGAGCGAGCTGAATTATCCCAACCGGTGGCAC 4631
Qy 181 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTATCCCAACCGGTGGCAC 240
Db 4630 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTATCCCAACCGGTGGCAC 4571
Qy 241 AACAACTGGCGGGCAACAGTCGTTGCTGATTGGCTTGCACCTCCAGTCTGCCCTGC 300
Db 4570 AACAACTGGCGGGCAACAGTCGTTGCTGATTGGCTTGCACCTCCAGTCTGCCCTGC 4511
Qy 301 ACGCCCGTCGCAAAATGTTCGGCGCGATTAATATCTCGCGCGATCAACTGGGTGCCAGCG 360

Db 4510 ACGCGCGTCGCAAAATGTTCGGCGCGATTAATATCTCGCGCGATCAACTGGGTGCCAGCG 4451
Qy 361 TGGTGGTGTGATGGTAGAAGCGCGCTGCAAGCTGTAAAGCGCGCGGTGCACAATC 420
Db 4450 TGGTGGTGTGATGGTAGAAGCGCGCTGCAAGCTGTAAAGCGCGCGGTGCACAATC 4391
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTGATGACACAGATGCCA 480
Db 4390 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTGATGACACAGATGCCA 4331
Qy 481 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGTGATGTTCTCTCACCAGA 540
Db 4330 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGTGATGTTCTCTGACCAGA 4271
Qy 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGCTAGCGACTGGGCGTGGAGATC 600
Db 4270 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGCTAGCGACTGGGCGTGGAGATC 4211
Qy 601 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAAGTTCTCTCTCGG 660
Db 4210 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAAGTTCTCTCTCGG 4151
Qy 661 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 720
Db 4150 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 4091
Qy 721 CGGAACGGGAAGCGACTGAGTGCCATGTCGGTTCGCGTTTGAACAAACCATGCAATGCTGA 780
Db 4090 CGGAACGGGAAGCGACTGAGTGCCATGTCGGTTCGCGTTTGAACAAACCATGCAATGCTGA 4031
Qy 781 ATGAGGSCATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 840
Db 4030 ATGAGGSCATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 3971
Qy 841 TGGCGGCATTCAGGATCGCGCTCGCGCTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 3970 TGGCGGCATTCAGGATCGCGCTCGCGCTGGTGGCGATATCTCGGTAGTGGGATACG 3911
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCAACATCAACACAGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCAACATCAACACAGATTTTC 3851
Qy 961 GCTCTGTTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 3850 GCTCTGTTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGA 3791
Qy 1021 AGGCAATCAGCTTGTGGCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
Db 3790 AGGCAATCAGCTTGTGGCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 3731
Qy 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGACAGTGGCACGACAGTTT 1140
Db 3730 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGACAGTGGCACGACAGTTT 3671
Qy 1141 CCGCATGGAACCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 1200
Db 3670 CCGCATGGAACCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 3611
Qy 1201 GCAC 1204
Db 3610 GCAC 3607

RESULT 11
US-10-263-103-35/c
; Sequence 35, Application US/10263103
; Patent No. 6821723
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PASTEUR
; APPLICANT: Chevalier, Michel
; APPLICANT: El Habib, Raphaelle
; APPLICANT: Krell, Tino
; APPLICANT: Sodoyer, Regis

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pET-cer
US-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGC	AAACCTTTCGGGTATGCGCATGATAGCCCGGAGAGA	60
DB	4622	CGGACACCATCGAATGGCG	CAAAACCTTTCGGGTATGCGCATGATAGCCCGGAGAGA	4563
QY	61	GTCGAATTCAGGGTGGTGA	ATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCCG	120
DB	4562	GTCGAATTCAGGGTGGTGA	ATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCCG	4503
QY	121	GTGTCTCTTATCAGACCG	TTTCCCGGTGGTGAACCAAGCCAGCCACGTTTCTGCGAAAA	180
DB	4502	GTGTCTCTTATCAGACCG	TTTCCCGGTGGTGAACCAAGCCAGCCACGTTTCTGCGAAAA	4443
QY	181	CGCGGAAAGTGAAGCGG	CGATGGCGAGCTGAATTACATTCGCAACCGGTGGCAC	240
DB	4442	CGCGGAAAGTGAAGCGG	CGATGGCGAGCTGAATTACATTCGCAACCGGTGGCAC	4383
QY	241	AACAACTCGCGGCAAA	CAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	4382	AACAACTCGCGGCAAA	CAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	4323
QY	301	ACGCGCCCTCGCAAA	TTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
DB	4322	ACGCGCCCTCGCAAA	TTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	4263
QY	361	TGGTGGTGTGATGTA	GAACGAAGCGCGCTGGAAGCCCTGTAAGCGCGGTGCAATC	420
DB	4262	TGGTGGTGTGATGTA	GAACGAAGCGCGCTGGAAGCCCTGTAAGCGCGGTGCAATC	4203
QY	421	TTCTCGCGCAACGCT	CAGTGGCTGATCAATTAATCTCCGTGGATGACCAAGATGCCA	480
DB	4202	TTCTCGCGCAACGCT	CAGTGGCTGATCAATTAATCTCCGTGGATGACCAAGATGCCA	4143
QY	481	TTGCTGTGGAAGCT	CGCTGCACATAATGTTTCGGCGGTTATTTCTGATGTCTGACCCAGA	540
DB	4142	TTGCTGTGGAAGCT	CGCTGCACATAATGTTTCGGCGGTTATTTCTGATGTCTGACCCAGA	4083
QY	541	CACCCATCAACGATTA	TATTTCTCCATGAAGCGGTACGGCACTGGGCGTGGAGCATC	600
DB	4082	CACCCATCAACGATTA	TATTTCTCCATGAAGCGGTACGGCACTGGGCGTGGAGCATC	4023
QY	601	TGGTCGCATGGGT	CACAGCAAAATCGCGCTGTAGCGGCCCATTAAGTTCTGTCGCG	660
DB	4022	TGGTCGCATGGGT	CACAGCAAAATCGCGCTGTAGCGGCCCATTAAGTTCTGTCGCG	3963
QY	661	CGCGTCTCGCT	GTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	3962	CGCGTCTCGCT	GTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG	3903
QY	721	CGGAAACGGGAAGCG	CACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA	780
DB	3902	CGGAAACGGGAAGCG	CACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA	3843
QY	781	ATGAGGGCATGTT	CCCCACTCGATGCTGGTTCACCAAGATCAGATGGCGTGGCGGCA	840
DB	3842	ATGAGGGCATGTT	CCCCACTCGATGCTGGTTCACCAAGATCAGATGGCGTGGCGGCA	3783

QY	841	TGCGCGCATTTACCGAGT	CCGGGCTCGCGCTTGGTGGGATATCTCGTAGTGGGATACG	900
DB	3782	TGCGCGCATTTACCGAGT	CCGGGCTCGCGCTTGGTGGGATATCTCGTAGTGGGATACG	3723
QY	901	ACGATACCGAAGACAGCT	CATGTTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC	960
DB	3722	ACGATACCGAAGACAGCT	CATGTTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC	3663
QY	961	GCCTGCTGGGCAAAAC	CAGCGTGGCACTCTCTCAGGGCCAGCGGTGA	1020
DB	3662	GCCTGCTGGGCAAAAC	CAGCGTGGCACTCTCTCAGGGCCAGCGGTGA	3603
QY	1021	AGGCAATCAGCTGTTG	CCCGCTCTCACTGCTGAAAGAAACCAACCTGGCGGCCAATA	1080
DB	3602	AGGCAATCAGCTGTTG	CCCGCTCTCACTGCTGAAAGAAACCAACCTGGCGGCCAATA	3543
QY	1081	CGCAAAACCGCTCT	CCCCCGCTTGGCGGATTCATTAATCAGTGGCAGCAGAGTTT	1140
DB	3542	CGCAAAACCGCTCT	CCCCCGCTTGGCGGATTCATTAATCAGTGGCAGCAGAGTTT	3483
QY	1141	CCGACTGGAAAGCGG	CAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG	1200
DB	3482	CCGACTGGAAAGCGG	CAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG	3423
QY	1201	GCAC	1204	
DB	3422	GCAC	3419	

RESULT 12
US-08-929-967-1/c
; Sequence 1, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Ruppert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGCGCATGATAGCGCCCGGAAGAGA 60
Db 4810 CGGACACCATCGAATGGGCGAAAACCTTTCCGGGTATGCGCATGATAGCGCCCGGAAGAGA 4751
QY 61 GTC AATTCAGGGTGGTGAATGTGAACCGATGAACCGTTATACGATGTCGAGAGTATGCCG 120
Db 4750 GTC AATTCAGGGTGGTGAATGTGAACCGATGAACCGTTATACGATGTCGAGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCGAGCCAGCCACGTTTCTCGGAAA 180
Db 4690 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCGAGCCAGCCACGTTTCTCGGAAA 4631
QY 181 CCGGGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db 4630 CCGGGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 4571
QY 241 AACAACTGCGGGGAAAACAGTCGTGCTGATTTGGGCTTTGCCACCTCCAGTCTGGCCCTGC 300
Db 4570 AACAACTGCGGGGAAAACAGTCGTGCTGATTTGGGCTTTGCCACCTCCAGTCTGGCCCTGC 4511
QY 301 ACGCGCCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTCTGGGTGCGACG 360
Db 4510 ACGCGCCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTCTGGGTGCGACG 4451
QY 361 TGGTGGTGCATGTGTAAGAACGAGCGCGTGAAGCGCTGAAGCGCGGTGACACAATC 420
Db 4450 TGGTGGTGCATGTGTAAGAACGAGCGCGTGAAGCGCGGTGACACAATC 4391
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACGACGATGCCA 480
Db 4390 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACGACGATGCCA 4331
QY 481 TTGCTGTGGAAGCTGCTCGACTAATGTTCCGGGCTTATTTCTTGATGCTCTCTCACAGA 540
Db 4330 TTGCTGTGGAAGCTGCTCGACTAATGTTCCGGGCTTATTTCTTGATGCTCTCTCACAGA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGCATCGGCGGTGAGCATC 600
Db 4270 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGCATCGGCGGTGAGCATC 4211
QY 601 TGGTCGCAATGGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTCTCTCGG 660
Db 4210 TGGTCGCAATGGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTCTCTCGG 4151
QY 661 CGCGTCTGCGTCTGGCTGGGCTAAATATCTCACTCGCAATCAAAATCAGCGCATAG 720
Db 4150 CGCGTCTGCGTCTGGCTGGGCTAAATATCTCACTCGCAATCAAAATCAGCGCATAG 4091
QY 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGGTTTCAACAAACCATGCAATGCTGA 780
Db 4090 CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGGTTTCAACAAACCATGCAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGTTGTCGAACGATCAGATGGCGCTGGGCGCAA 840
Db 4030 ATGAGGGCATCGTTCCCACTGCGATGCTGTTGTCGAACGATCAGATGGCGCTGGGCGCAA 3971
QY 841 TGCGCGCATTAACGAGTCCGGGCTCGCGTGGTGGTGCGATATCTCGGTAGTGGGATACG 900
Db 3970 TGCGCGCATTAACGAGTCCGGGCTCGCGTGGTGGTGCGATATCTCGGTAGTGGGATACG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAACAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAACAGGATTTTC 3851
QY 961 GCTGTGGGGGAAAACAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 3850 GCTGTGGGGGAAAACAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 3791
QY 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGTTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 3790 AGGCAATCAGCTGTTGCCGCTCTCACTGTTGAAAGAAAAACCAACCTGGCGCCCAATA 3731
QY 1081 CGAAACCGGCTCTCTCCCGCGCGTGGCGGATTCATTAATGCGAGTGGCACGACAGTTT 1140
```

```
Db 3730 CGAAACCGGCTCTCTCCCGCGCGTGGCGGATTCATTAATGCGAGTGGCACGACAGTTT 3671
QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGCTTAGCTCACTCATTAG 1200
Db 3670 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGCTTAGCTCACTCATTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607

RESULT 13
US-09-702-705-785/c
; Sequence 785, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCES: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGCGCATGATAGCGCCCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGGGCGAAAACCTTTCCGGGTATGCGCATGATAGCGCCCGGAAGAGA 4618
QY 61 GTC AATTCAGGGTGGTGAATGTGAACCGATGAACCGTTATACGATGTCGACAGATATGCCG 120
Db 4617 GTC AATTCAGGGTGGTGAATGTGAACCGATGAACCGTTATACGATGTCGACAGATATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCGAGCTGAATTAATCCCAACCGCGTGGCAC 180
Db 4557 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCGAGCTGAATTAATCCCAACCGCGTGGCAC 4498
QY 181 CCGGGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db 4497 CCGGGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 4438
QY 241 AACAACTGCGGGGAAAACAGTCGTGCTGATTTGGGCTTTGCCACCTCCAGTCTGGCCCTGC 300
Db 4437 AACAACTGCGGGGAAAACAGTCGTGCTGATTTGGGCTTTGCCACCTCCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGCTCGAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTCAACAGGCGGTGCAAAATC 360
Db 4377 ACGCGCGCTCGAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTCAACAGGCGGTGCAAAATC 4318
QY 361 TGGTGGTGCATGTGTAAGAACGAGCGGTGCAAGCGCTGTAAAGCGGCGGTGCAAAATC 420
Db 4317 TGGTGGTGCATGTGTAAGAACGAGCGGTGCAAGCGCTGTAAAGCGGCGGTGCAAAATC 4258
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCCGCTGGATGACGAGGATGCCA 480
```

Db 4257 TTCTCGCAACGCGTCAGTGGCTGATCATTAACATATCCGCTGGATGACCAAGATGCCA 4198
Qy 481 TTGCTGTGGAAGTCCCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 540
Db 4197 TTGCTGTGGAAGTCCCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 4138
Qy 541 CACCCATCAACAGTATATTTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATATTTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATC 4078
Qy 601 TGGTGCATTTGGGTCCACGACAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCCTGG 660
Db 4077 TGGTGCATTTGGGTCCACGACAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCGG 4018
Qy 661 CGCGTCTGCGTCTGCGTGGCTGCGATTAATATCTCACTCGGATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCTGCGATTAATATCTCACTCGGATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAATGCTGA 3898
Qy 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGCATGAGTGGCGCGCAA 840
Db 3897 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGCATGAGTGGCGCGCAA 3838
Qy 841 TGCAGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 3837 TGCAGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 3778
Qy 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAACAGGATTTTC 960
Db 3777 ACGATACGGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAACAGGATTTTC 3718
Qy 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 3658
Qy 1021 AGGCAATCAGCTGTGCGCGTCTCACTGTTGAAAGAAACCAACCTGGCGCGCAATA 1080
Db 3657 AGGCAATCAGCTGTGCGCGTCTCACTGTTGAAAGAAACCAACCTGGCGCGCAATA 3598
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTATGCACTGSCAGCTGSCAGAGTTT 1140
Db 3597 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTATGCACTGSCAGCTGSCAGAGTTT 3538
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 3537 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 3478
Qy 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 14

US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Panger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGGCAAAACCTTTTCGGGTATGGCATGATAGGCCCGGGAAGAGA 60
Db 4677 CCGACACCATCGAATGGCGCAAAACCTTTTCGGGTATGGCATGATAGGCCCGGGAAGAGA 4618
Qy 61 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCG 4558
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCAGCCAGCTTTCTGCAAAA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCAGCTTTCTGCAAAA 4498
Qy 181 CGCGGAAAAAGTGGAGCGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGGAGCGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 4438
Qy 241 AACAACTGGCGGGCAACAGTCTGTTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAACAGTCTGTTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 4378
Qy 301 ACGCGCGTCCGCAAAATTCCTCGCGCGATTAATCTCGCGCGATCAACTGGGTGGCAGCG 360
Db 4377 ACGCGCGTCCGCAAAATTCCTCGCGCGATTAATCTCGCGCGATCAACTGGGTGGCAGCG 4318
Qy 361 TGGTGGTTCGATGGTGAACAGCGCGCTGCAAGCCTGTAAGCGCGGTGCAACAATC 420
Db 4317 TGGTGGTTCGATGGTGAACAGCGCGCTGCAAGCCTGTAAGCGCGGTGCAACAATC 4258
Qy 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATATCCCTGGATGACCAAGGATGCA 480
Db 4257 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATATCCCTGGATGACCAAGGATGCA 4198
Qy 481 TTGTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 540
Db 4197 TTGTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA 4138
Qy 541 CACCCATCAACAGTATTAATTTTCTCCCATGAAGACGGTACGCACTGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTAATTTTCTCCCATGAAGACGGTACGCACTGGCGGTGGAGCATC 4078
Qy 601 TGGTGCATTTGGGTCCACGACAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCGG 660
Db 4077 TGGTGCATTTGGGTCCACGACAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCGG 4018
Qy 661 CGCGTCTGCGTCTGCGTGGCTGCGATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCTGCGATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAATGCTGA 3898
Qy 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGCATGAGTGGCGGTGGCGCAA 840
Db 3897 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGCATGAGTGGCGGTGGCGCAA 3838
Qy 841 TGCAGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 3837 TGCAGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 3778

```
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACACCATCAAAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACACCATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 3658
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCCCTGGGGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCCCTGGGGCCCAATA 3598
QY 1081 CGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGACGACAGGTTT 1140
Db 3597 CGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGACGACAGGTTT 3538
QY 1141 CCCGACTGGAAAGCGGGCAGTGAGCGCAACCGAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 3537 CCCGACTGGAAAGCGGGCAGTGAGCGCAACCGAATTAATGTGAGTTAGCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 15
US-09-614-124B-785/c
; Sequence 785, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedrick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614.124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGTGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 4618
QY 61 GTCAATTCAGGTTGTTGAATGTGAACAGTAAAGTTATAGATGTGCGAGATATGCCG 120
Db 4617 GTCAATTCAGGTTGTTGAATGTGAACAGTAAAGTTATAGATGTGCGAGATATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGTGTAACACGAGCCAGCGACGTTCTTTCGCAAAA 180
Db 4557 GTGTCTCTTATCAGACCGGTTTCCCGCGTGTGTAACACGAGCCAGCGACGTTCTTTCGCAAAA 4498
QY 181 CGCGGAAAAAGTGGAAAGCGGCGATGGCGAGCTGGAATTAATTCCAAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGGAAAGCGGCGATGGCGAGCTGGAATTAATTCCAAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAACAGCTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
```

Search completed: August 3, 2005, 02:32:31.
Job time : 356.4 secs

```
Db 4437 AACAACTGGCGGGCAACAGCTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGTCCGCAAAATTTGTTCGCGCGATTAATCTCGCGCGATCAACTCGGTGCCAGCG 360
Db 4377 ACGCGCGTCCGCAAAATTTGTTCGCGCGATTAATCTCGCGCGATCAACTCGGTGCCAGCG 4318
QY 361 TGGTGTGTGNTGGTGAAGCGAGCGCGTTCGAAAGCTGTAAAGCGCGGTGCACAATC 420
Db 4317 TGGTGTGTGNTGGTGAAGCGAGCGCGTTCGAAAGCTGTAAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACAGGATGCCA 480
Db 4257 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACAGGATGCCA 4198
QY 481 TTGCTGTGGAAGCTGCCTGCACATAAATTCGCGGTTTATTTCTTGATGTCTCTGACCAGA 540
Db 4197 TTGCTGTGGAAGCTGCCTGCACATAAATTCGCGGTTTATTTCTTGATGTCTCTGACCAGA 4138
QY 541 CACCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 4137 CACCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 4078
QY 601 TGGTTCGCAATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCGG 660
Db 4077 TGGTTCGCAATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCGG 4018
QY 661 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTGCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTGCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGAAACGGGAAGCGACTGGAGTGCATGTCGGTTTTCACCAACCATCAATGCTGA 780
Db 3957 CGAAACGGGAAGCGACTGGAGTGCATGTCGGTTTTCACCAACCATCAATGCTGA 3898
QY 781 ATGAGGCGATCGTTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGGGCGCAA 840
Db 3897 ATGAGGCGATCGTTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGGGCGCAA 3838
QY 841 TGGCGCGCAATTCAGGATTCGGGCTCGGCTGGGCTGGGATATCTCGGTAGTGGGATACG 900
Db 3837 TGGCGCGCAATTCAGGATTCGGGCTCGGCTGGGCTGGGATATCTCGGTAGTGGGATACG 3778
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTTAAACACCATCAAAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTTAAACACCATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 3658
QY 1021 AGGGCAATCAGCTGTTGGCGGCTCTCACTGGTGAAGAAACCAACCCCTGGGGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTGGCGGCTCTCACTGGTGAAGAAACCAACCCCTGGGGCCCAATA 3598
QY 1081 CGAAACCGCTCTCCCGCGGCTTCGCGGATTCATTAAATGAGCTGGCAGACGACAGGTTT 1140
Db 3597 CGAAACCGCTCTCCCGCGGCTTCGCGGATTCATTAAATGAGCTGGCAGACGACAGGTTT 3538
QY 1141 CCGCACTGGAAAGCGGGCAGTGAGCGCAACCGAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 3537 CCGCACTGGAAAGCGGGCAGTGAGCGCAACCGAATTAATGTGAGTTAGCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds

(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-17

Perfect score: 3300

Sequence: 1 ccgacacatcgatggtgc.....acgacgttcggagtaagct 3300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3300	100.0	3300	10	US-09-765-555-17
2	3232.8	98.0	3300	10	US-09-765-555-18
3	3215.2	97.4	3300	10	US-09-765-555-14
4	3215.2	97.4	3300	10	US-09-765-555-15
5	3207.2	97.2	3300	10	US-09-765-555-16
6	2705.8	82.0	7373	18	US-09-765-384A-1
7	2705.8	82.0	8101	17	US-10-149-472-5

8	2701	81.8	6648	19	US-10-343-859-8	Sequence 8, Appli
9	2698	81.8	9191	19	US-10-343-859-9	Sequence 9, Appli
10	2690.4	81.5	7112	18	US-10-263-153-40	Sequence 40, Appli
11	2690.4	81.5	7259	18	US-10-263-153-30	Sequence 30, Appli
12	2690.4	81.5	7322	18	US-10-263-153-35	Sequence 35, Appli
13	2690.4	81.5	7352	18	US-10-263-153-25	Sequence 25, Appli
14	2690.4	81.5	7370	18	US-10-263-153-20	Sequence 20, Appli
15	2690.4	81.5	7370	18	US-10-263-153-61	Sequence 61, Appli
16	2690.4	81.5	7370	18	US-10-263-153-66	Sequence 66, Appli
17	2690.4	81.5	7370	18	US-10-263-153-71	Sequence 71, Appli
18	2690.4	81.5	7403	18	US-10-263-153-15	Sequence 15, Appli
19	2690.4	81.5	7442	18	US-10-263-153-10	Sequence 10, Appli
20	2690.4	81.5	7478	18	US-10-263-153-3	Sequence 3, Appli
21	2636.8	79.9	6806	10	US-09-908-943A-194	Sequence 194, App
22	2636.8	79.9	6806	20	US-10-801-487-194	Sequence 194, App
23	2636.8	79.9	6806	20	US-10-801-938-194	Sequence 194, App
24	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
25	2636.8	79.9	6806	20	US-10-801-486-194	Sequence 194, App
26	2636.8	79.9	6806	21	US-10-801-493-194	Sequence 194, App
27	2605.4	79.0	7553	18	US-10-263-153-7	Sequence 7, Appli
28	1609.4	48.8	5558	16	US-10-241-596-137	Sequence 137, App
29	1474	44.7	4700	17	US-10-313-963A-54	Sequence 54, Appli
30	1462.4	44.3	4921	21	US-10-495-491-2	Sequence 2, Appli
31	1462.4	44.3	4935	21	US-10-495-491-1	Sequence 1, Appli
32	1462.4	44.3	4945	21	US-10-495-491-3	Sequence 3, Appli
33	1462.4	44.3	4951	21	US-10-495-491-4	Sequence 4, Appli
34	1417.6	43.0	1511	10	US-09-987-763-30	Sequence 30, Appli
35	1417.6	43.0	5926	13	US-10-024-809-3	Sequence 3, Appli
36	1238.4	37.5	5024	17	US-10-313-963A-52	Sequence 52, Appli
37	1238.4	37.5	5448	21	US-10-671-995A-2	Sequence 2, Appli
38	1238.4	37.5	5454	21	US-10-671-995A-1	Sequence 1, Appli
39	1237.2	37.5	1922	10	US-09-995-988A-32	Sequence 32, Appli
40	1237.2	37.5	1922	10	US-10-420-034A-32	Sequence 32, Appli
41	1237.2	37.5	6270	21	US-10-855-897-2	Sequence 2, Appli
42	1204	36.5	4557	17	US-10-457-372-5	GENERAL INFORMA
43	1204	36.5	5664	18	US-10-270-176-1	Sequence 1, Appli
44	1204	36.5	5668	18	US-10-270-176-2	Sequence 2, Appli
45	1204	36.5	6742	18	US-10-270-176-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-09-765-555-17
; Sequence 17, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein 2Fpm4
US-09-765-555-17

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTCGCAAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTCGCAAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACTGATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACTGATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCCTTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCACGTTTCTGGGAAA 180
Db 121 GTGTCCTTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCACGTTTCTGGGAAA 180
QY 181 CCGGGGAAAAAGTGAAGCGGCGATGCGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CCGGGGAAAAAGTGAAGCGGCGATGCGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACCACTGGCGGCAAAACAGTCTGTCATTTGGGCTTGGCCACCTTCCAGTCTGGCCCTGC 300
Db 241 AACCACTGGCGGCAAAACAGTCTGTCATTTGGGCTTGGCCACCTTCCAGTCTGGCCCTGC 300
QY 301 AGCGCCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCCGATCAAATGGGTCGACGCG 360
Db 301 AGCGCCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCCGATCAAATGGGTCGACGCG 360
QY 361 TGGTGTGTCGATGCTGAGAACGAAGCGGCTGAAAGCCCTGTAAGCGGCGGTGCAATC 420
Db 361 TGGTGTGTCGATGCTGAGAACGAAGCGGCTGAAAGCCCTGTAAGCGGCGGTGCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATCACCAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATCACCAGGATGCCA 480
QY 481 TTGCTGTGAAGCTGCTGCAATAATGTCGGCGGTTATTTCTTGATGTCCTGACACAGA 540
Db 481 TTGCTGTGAAGCTGCTGCAATAATGTCGGCGGTTATTTCTTGATGTCCTGACACAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAGAGGTTACGGACTGGCGGTGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAGAGGTTACGGACTGGCGGTGAGCATC 600
QY 601 TGGTCGCAATGGTCCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCGCG 660
Db 601 TGGTCGCAATGGTCCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCGCG 660
QY 661 CGCGTCTGGCTGGCTGGCTGACATAATCTCACTCGCAATCAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTGGCTGGCTGACATAATCTCACTCGCAATCAAATTCAGCCGATAG 720
QY 721 CCGAACCGGAAGGCGACTGGAGTGCATGTCGGTTCGCAACCATGCAATGCTGA 780
Db 721 CCGAACCGGAAGGCGACTGGAGTGCATGTCGGTTCGCAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGATGTCGATGTCGCAACCATGCAATGCTGA 840
Db 781 ATGAGGGCATCGTTCCCACTCGATGTCGATGTCGCAACCATGCAATGCTGA 840
QY 841 TGGCGCCATTAACAGTTCGGGCTGGCGTGGTGGCGATATCTCGTATGCGGATACG 900
Db 841 TGGCGCCATTAACAGTTCGGGCTGGCGTGGTGGCGATATCTCGTATGCGGATACG 900
QY 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAACAGGATTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGAACCGCTTGTGCAACTCTCTCAGGCGCAAGCGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGAACCGCTTGTGCAACTCTCTCAGGCGCAAGCGTGA 1020
QY 1021 AGGCAATACAGCTTGGCCGCTCTCACTGTGAAAGAAAAACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATACAGCTTGGCCGCTCTCACTGTGAAAGAAAAACCCCTGGCGCCCAATA 1080
QY 1081 CGMAACCGGCTCTCCCCGGCGGTTGGCGATTCATTTAATGTCAGCTGGCACGACAGGTTT 1140

Db 1081 CGMAACCGGCTCTCCCCGGCGGTTGGCGATTCATTTAATGTCAGCTGGCACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCAATTAATGTAGTGTAGCTCACCTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCAATTAATGTAGTGTAGCTCACCTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCAACGCTGCAACAAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCAACGCTGCAACAAATGCTTCTGGCG 1260
QY 1261 TCAGGCAAGCCATCGGAAAGCTGTGATGCTGTGCTGCAAGCTGTAATCACTGCAATAATTCG 1320
Db 1261 TCAGGCAAGCCATCGGAAAGCTGTGATGCTGTGCTGCAAGCTGTAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTTGGCCGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTTGGCCGACATCATAAACGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
QY 1441 ATTGTGAGCGGATAAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAAAGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAAAGTGTTTTCACGA 1500
QY 1501 GCATTTCAACAAACAGGACCATAGATTAATGAAACTGAAGAGGTAATCTGGTAAATCTCG 1560
Db 1501 GCATTTCAACAAACAGGACCATAGATTAATGAAACTGAAGAGGTAATCTGGTAAATCTCG 1560
QY 1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT 1680
QY 1681 GCGGCAACTGGCCATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGCTAC 1740
Db 1681 GCGGCAACTGGCCATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGACAAAGCTTCCAGACAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGACAAAGCTTCCAGACAGCTGAT 1800
QY 1801 CGTTTTACCTGGGATGCCGTACGTTACAAACGCAAGCTGATTTACCCGATCGCTGTT 1860
Db 1801 CGTTTTACCTGGGATGCCGTACGTTACAAACGCAAGCTGATTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTTCAAC 1980
QY 1981 CTGCAAGAACCTTACCTTACCTGGCGCTGATGCTGCTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCTTACCTTACCTGGCGCTGATGCTGCTGACGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGAACCTTCTGCTGTTGAACCTGATTAATAAACAATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGAACCTTCTGCTGTTGAACCTGATTAATAAACAATGAATGACACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220

Db 601 |||||TGTCGCATTTGGGTCAACGAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CCGGTCTGGCTCTGGCTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CCGGTCTGGCTCTGGCTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CCGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAACGGGAAGCGGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCTCCCATCGGATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCTCCCATCGGATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGAA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGCGATATCTCGTGTGGGATACG 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGCGATATCTCGTGTGGGATACG 900
QY 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
Db 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGTGGACGCTTGCCTGCACTCTCTCAGGCGCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACGCTTGCCTGCACTCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACCAACCTCGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACCAACCTCGCGGCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGTAGCTCAGTGGCAGCAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGTAGCTCAGTGGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGAGTGGCGCAACGCAATTAATGAGTGTAGCTCAGTGGCAGCAGGTTT 1200
Db 1141 CCCGACTGGAAGCGGAGTGGCGCAACGCAATTAATGAGTGTAGCTCAGTGGCAGCAGGTTT 1200
QY 1201 GCACAAATTCATGTTGACAGCTTATCATCGACTGACCGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTCATGTTGACAGCTTATCATCGACTGACCGTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAGCTGTGCTATGGCTGTGCGAGTGTGAATCACTGCAATAATTCG 1320
Db 1261 TCAGCAGCCATCGGAGCTGTGCTATGGCTGTGCGAGTGTGAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCATACCGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTGGTTTACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTGGTTTACGA 1500
QY 1501 GCATTTCAACCAAGGACCATAGATTATGAAATCTGAAGAGTAACTGGTAATCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATTATGAAATCTGAAGAGTAACTGGTAATCTGG 1560
QY 1561 ATTTAACCGGATAAAGGCTATACGGTCTCGTGAAGTCCGGTAAAGAAATTCGAGAAGAT 1620
Db 1561 ATTTAACCGGATAAAGGCTATACGGTCTCGTGAAGTCCGGTAAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 CGCGCAATGGCGCTCGCATTAATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740

Db 1681 CGCGCAATGGCGCTCGCATTAATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGCAAAAGGTTTCCAGGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGCAAAAGGTTTCCAGGCAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGTGCGCAACCCGCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGTGCGCAACCCGCCAAAACCTCGGAA 1920
QY 1921 GAGATCCCGGCGTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGTATCAAC 1980
Db 1921 GAGATCCCGGCGTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGTATCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGCTGACGGGTTTATGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGCTGACGGGTTTATGGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGGCGTGCATTAACGCTGCGCGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGGCGTGCATTAACGCTGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGACCTGATTTAAACAAACACATGATGAGATCGGATTAAC 2160
Db 2101 GGTCTGACCTTCTGCTGACCTGATTTAAACAAACACATGATGAGATCGGATTAAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
QY 2221 GCATGTCCAAATCAGACACAGCAAAAGTGAATTTATGGTGAACGCTGATCGCGACCTTC 2280
Db 2221 GCATGTCCAAATCAGACACAGCAAAAGTGAATTTATGGTGAACGCTGATCGCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGAGCGCAGGATTTAACGCCGCACT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGAGCGCAGGATTTAACGCCGCACT 2340
QY 2341 CCGAAACAAAGAGCTGCGAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGCGAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGCTCGGTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGCTCGGTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGCTCAGACTGCTGATGAAGCCCTGGAAGCGCCGACACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTCAGACTGCTGATGAAGCCCTGGAAGCGCCGACACTTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACAACTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAAACAAATAACAACTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
QY 2701 GATCTCTTCTCTGTGGCCAGCGGCGCTTCTGAGCCCGGGGGAAGGCCCTTATGCTTGT 2760
Db 2701 GATCTCTTCTCTGTGGCCAGCGGCGCTTCTGAGCCCGGGGGAAGGCCCTTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTTTCAGCGCAGAGCAGCTCCCTGGTGGCGCACCCAGGTAACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTTCAGCGCAGAGCAGCTCCCTGGTGGCGCACCCAGGTAACCCAC 2820

```
QY 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTCGGCGCAAAATCTTTTAGCCAGAGCAGAGC 2880
Db 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTCGGCGCAAAATCTTTTAGCCAGAGCAGAGC 2880
QY 2881 CTGGTGGCCATCAACGACCTACTGCGGAGAGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGACCTACTGCGGAGAGCCATACAAATGTCCAGATGTGGC 2940
QY 2941 AAGTCTTTTCAGTGATGTGCGTATCTTGGAGGACCAACGCTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTTCAGTGATGTGCGTATCTTGGAGGACCAACGCTACTCACACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCACCAGTGGCTCTCTGGTTAGACAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTCGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTCGGCGCAAAATCTTTTAGC 3120
QY 3121 CGACGCGTAACTGCTGGCCATCAACGCACTCATCTACTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGCCCACTTGAAGCCATCAACGCACTCATCTACTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGATGTGGCAAGTCTTTCTCAACTTCAGGCCATTTGGTCCGTCAACCAAGTACTCAC 3240
Db 3181 CCAGATGTGGCAAGTCTTTCTCAACTTCAGGCCAATTTGGTCCGTCAACCAAGTACTCAC 3240
QY 3241 ACCGCTAAAAAAGTGTGGCCAGCGCGCCAGTACCGGTACGAGCTTCGCGACTACGCT 3300
Db 3241 ACCGCTAAAAAAGTGTGGCCAGCGCGCCAGTACCGGTACGAGCTTCGCGACTACGCT 3300
```

RESULT 3

```
US-09-765-555-14
; Sequence 14, Application US/0976555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m1 and zinc finger
; OTHER INFORMATION: protein ZFPm1
US-09-765-555-14
```

```
Query Match 97.4%; Score 3215.2; DB 10; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 CCGACACCATCAATGGTGCAGAAACCTTTCGGGTATGCGATGATAGCGCCGGAGAGA 60
Db 1 CCGACACCATCAATGGTGCAGAAACCTTTCGGGTATGCGATGATAGCGCCGGAGAGA 60
QY 61 GTCAATTCAGGTTGGTGAATGTGAACACAGTAACTTATACCATGTCGAGAGTATGCG 120
Db 61 GTCAATTCAGGTTGGTGAATGTGAACACAGTAACTTATACCATGTCGAGAGTATGCG 120
QY 121 GTGTCCTTATCAGACCGGTTTCCCGGTGGTGAACACGAGCCAGCCAGCTTTCTGCGAAA 180
```

```
Db 121 GTGTCCTTATCAGACCGGTTTCCCGGTGGTGAACACGAGCCAGCCAGCCAGCTTTCTGCGAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGGCAACAGTCGTTGCTGATGTGGGTTGGCCACTCTCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGGCAACAGTCGTTGCTGATGTGGGTTGGCCACTCTCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATGTTCGCGCGATTAATCTCGCGCGATCAACTCGGTGTCGACG 360
Db 301 ACGCGCGTTCGCAAAATGTTCGCGCGATTAATCTCGCGCGATCAACTCGGTGTCGACG 360
QY 361 TGGTGGTGTGATGATAGAACGAAAGCGCGTGAAGCGCTGTAAGCGCGGTGACCAATC 420
Db 361 TGGTGGTGTGATGATAGAACGAAAGCGCGTGAAGCGCGGTGACCAATC 420
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATATATATATCCGCTGGATGACCAAGTATGCA 480
Db 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATATATATATCCGCTGGATGACCAAGTATGCA 480
QY 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCGCGGCTTAATTTCTTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCGCGGCTTAATTTCTTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGAGCGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCAATGGGTGTCACAGCAAAATCGCGCTGTGTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db 601 TGGTCGCAATGGGTGTCACAGCAAAATCGCGCTGTGTAGCGGGCCCATTAAGTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCTGTTCCCACTCGCATGTGTGTGCAACAGATCAGATGCGGTGGCGCGAA 840
Db 781 ATGAGGGCATCTGTTCCCACTCGCATGTGTGTGCAACAGATCAGATGCGGTGGCGCGAA 840
QY 841 TGGCGCCATTACGAGTTCGGGCTGGCGGTGGTGGGATATCTCGTGTGGGATAGC 900
Db 841 TGGCGCCATTACGAGTTCGGGCTGGCGGTGGTGGGATATCTCGTGTGGGATAGC 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGGG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGGG 1260
```

Db 1201 GCACAATTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGCAGCCATCGGAAGCTGTGGTATATGCTGTGAGGTGTAATCACTGATCAATTCG 1320
Db 1261 TCAGGCAGCCATCGGAAGCTGTGGTATATGCTGTGAGGTGTAATCACTGATCAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
Qy 1381 CTGGCAATATCTCAATAGAGCTGTGCACATTAATCATCGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATCTCAATAGAGCTGTGCACATTAATCATCGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTGTTTTCAGGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTGTTTTCAGGA 1500
Qy 1501 GCACCTTCCAAACAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGTTAATCTGG 1560
Db 1501 GCACCTTCCAAACAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGTTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT 1680
Qy 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCGGACAAAGGTTCCAGGACAACTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCGGACAAAGGTTCCAGGACAACTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGTACGTTTACAAACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGTACGTTTACAAACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGTGCGGAACCCGCGCAAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGTGCGGAACCCGCGCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGATTAAGACGTTGAGCGTGAATGAGACACCGGATTAC 2100
Db 2041 TATGAAACCGCAAGTACGATTAAGACGTTGAGCGTGAATGAGACACCGGATTAC 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGATGAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGATGAGACACCGGATTAC 2160
Qy 2161 TCCATCGCAGNAGTGCCTTTTAAAGGCGAAACAGCGATGACCATCAACCGGCCGTGG 2220
Db 2161 TCCATCGCAGNAGTGCCTTTTAAAGGCGAAACAGCGATGACCATCAACCGGCCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACAGCAAGTGAATTTATGTTGTAAGGCTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACAGCAAGTGAATTTATGTTGTAAGGCTGCGGACCTTC 2280
Qy 2281 AAGGTTCAACATCCAAACCGTTGCTGTTGGCTGCTGAGCGCAGGTATTAACCGCCGCACT 2340
Db 2281 AAGGTTCAACATCCAAACCGTTGCTGTTGGCTGCTGAGCGCAGGTATTAACCGCCGCACT 2340

Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGTGACTGATGAAGGCTGT 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGTGACTGATGAAGGCTGT 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCTGAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCTGAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGCGAAAAGATCCACGTATTGCGCGCACCATGGAACCCAGAAAAGGTGAATCATG 2520
Db 2461 TTGCGAAAAGATCCACGTATTGCGCGCACCATGGAACCCAGAAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGATGCGGTGCTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGAGACTGTGATGAAGCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGAGACTGTGATGAAGCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATAAACAACAACTCGGATCGAGGAGGATTTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAAACAACAACTCGGATCGAGGAGGATTTTTCAGAAATTC 2700
Qy 2701 GGATCTCTTCTCTGTGGCCAGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGGTAAAGTCTTTTCAGCCAGAGCAGCTCCCTGCTGGTGGCCACACAGCGTACCAC 2820
Db 2761 CCGGAATGTGGTAAAGTCTTTCTCTCAGAGCTCTCACTGTGTGGCCACACAGCGTACCAC 2820
Qy 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTATGCCAGAGCAGCAGC 2880
Db 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTATGCCAGAGCAGCAGC 2880
Qy 2881 CTGCTGGCCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGCTGGCCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Qy 2941 AAGTCTTTTCAAGTGTGCTGATCTTTCAGAGAGCACAACAGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTTCTCTCGGTCTGACAACTCTCGTCCGCAACCAAGTACTCACACCGGGGAGAAG 3000
Qy 3001 CCTATGCTGTGTCGGGAATGTGTAAGTCTTCTCTCAGAGTCTCTCACCTGTGTGCCAC 3060
Db 3001 CCTATGCTGTGTCGGGAATGTGTAAGTCTTCTCAGCGCAGCGATAACTGTGTGCCAC 3060
Qy 3061 CAGCGTACCCACACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTTAGC 3120
Db 3061 CAGCGTACCCACACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTTAGC 3120
Qy 3121 CGCAGCGATAACTCTGTGCGCCATCAACGCACTCATATACTGGGAGAGGCAATACAATGT 3180
Db 3121 CAGCGCGGCGCACCTTGCCAGCCATCAACGCACTCATATACTGGGAGAGGCAATACAATGT 3180
Qy 3181 CCAGAATGTGGCAAGTCTTTCTCAACTTCAGGCCATTTGCTCGGTCCACCACTACTCAC 3240
Db 3181 CCAGAATGTGGCAAGTCTTTCTCGGTCTGCAATCTCTCGTCCGCGCACCAAGCTACTCAC 3240
Qy 3241 ACCGGTAAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTAGCAGCTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTAGCAGCTTCCGGACTACGCT 3300

RESULT 4

US-09-765-555-15
; Sequence 15, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate

; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US 09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger
; OTHER INFORMATION: protein ZFPm2
US-09-765-555-15

Query Match 97.4%; Score 3215.2; DB 10; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 CGACACCATGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CGACACCATGAAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60

QY 61 GTCAAATTCAGGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG 120
DB 61 GTCAAATTCAGGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG 120

QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGCGTTTCTCGGAAAA 180
DB 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGCGTTTCTCGGAAAA 180

QY 181 CGCGGAAAAAGTGAAGCGGGGATGGCGAGCTCAATTTACATTCCTCAACCGCGTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGGGGATGGCGAGCTCAATTTACATTCCTCAACCGCGTGGCAC 240

QY 241 AACAACTGGCGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

QY 301 AGCGCCGTGCGAAATTTGTCGGCGGATTAATCTCGCGCGGATCAATCTGGGTGCCAGCG 360
DB 301 AGCGCCGTGCGAAATTTGTCGGCGGATTAATCTCGCGCGGATCAATCTGGGTGCCAGCG 360

QY 361 TGGTGGTTCGATGCTAGAACGAGCGGTCGAGCCCTGTAAAGCGGGTGCACAATC 420
DB 361 TGGTGGTTCGATGCTAGAACGAGCGGTCGAGCCCTGTAAAGCGGGTGCACAATC 420

QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480

QY 481 TTGCTGTGGAAGCTGCCCTGCATTAATGTTCCGGCGTATTTCTTGATGCTCTGACACAGA 540
DB 481 TTGCTGTGGAAGCTGCCCTGCATTAATGTTCCGGCGTATTTCTTGATGCTCTGACACAGA 540

QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600

QY 601 TGGTGGCATTTGGTTCACAGCAAAATCGCGCTGTATAGCGGGCCCAATTAAGTTCTGTCGG 660
DB 601 TGGTGGCATTTGGTTCACAGCAAAATCGCGCTGTATAGCGGGCCCAATTAAGTTCTGTCGG 660

QY 661 CGCGTCTGCGTCTGCGTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGCGTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780

DB 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCTGTTCCCATCTCGCATGTGGTTGGCAACAGATCAGATGGCGCTGGCGCA 840
DB 781 ATGAGGGCATCTGTTCCCATCTCGCATGTGGTTGGCAACAGATCAGATGGCGCTGGCGCA 840
QY 841 TGGCGGCATTTACAGATGTCGGGCTGGCGGTGGTGGCGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGGCATTTACAGATGTCGGGCTGGCGGTGGTGGCGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGGTTTAAACCAACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGGTTTAAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAACCAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGGCAACCAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCTCTCACTGTGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCTCTCACTGTGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCAGCTGGCAGCAGCAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCAGCTGGCAGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTCAGTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTCAGTCACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTGACAGCTTATCATCGATGCAAGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATTCATGTTTGACAGCTTATCATCGATGCAAGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATAAATTCG 1320
DB 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCATACCGGTT 1380
DB 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCATACCGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGTAGTTCACGA 1500
DB 1441 ATTTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGTAGTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGGAAGAGGTAAACTGGTAACTCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGGAAGAGGTAAACTGGTAACTCTGG 1560
QY 1561 ATTTAAGCGGATTAAGGCTATAAAGCTCTCGCTGAGTTCGGTGAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTTAAGCGGATTAAGGCTATAAAGCTCTCGCTGAGTTCGGTGAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGAGAAAGCGTTCCAGGACAGCTGTAT 1800
DB 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGAGAAAGCGTTCCAGGACAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCGAGCTGATTGCTTACCCGATCCGTT 1860
DB 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCGAGCTGATTGCTTACCCGATCCGTT 1860


```
QY 2461 TTGGCGAAGATCCAGCTATTGCGCCGACCATGGAAAAAGGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCAGCTATTGCGCCGACCATGGAAAAAGGCGCCAGAAAGGTGAATCATG 2520
QY 2521 CGGAACATCCGCGACATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
Db 2521 CGGAACATCCGCGACATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTGATGAAGCCCTTGAAGAGCGCGACACTAAATTCAGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTGATGAAGCCCTTGAAGAGCGCGACACTAAATTCAGAGCTCG 2640
QY 2641 AACACACACACATAAACAATAAACAACCTCGGGATCGAGGAGGAAGATTTCAGAAATTC 2700
Db 2641 AACACACACACATAAACAATAAACAACCTCGGGATCGAGGAGGAAGATTTCAGAAATTC 2700
QY 2701 GGATCCTCTCTCTGTCGCGCCAGCGCCCTCGAGCGCCGCGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCCTCTCTCTGTCGCGCCAGCGCCCTCGAGCGCCGCGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTTTCAGCCAGAGAGCTCCCTGGTGGCCACACGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTTCAGCCAGATCTTTCAGCCACCTGGTTCGCCACGAGTACCCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTAGCCAGAGCAGCAGC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTAGCCAGCAGCAGCAGC 2880
QY 2881 CTGGTGGCCCATCAACGACCTCATCTGGCGAGAGCCATACAAATGTCAGAGATGGGC 2940
Db 2881 CTGGTGGCCCATCAACGACCTCATCTGGCGAGAGCCATACAAATGTCAGAGATGGGC 2940
QY 2941 AAGTCTTTCACTGATGTCGTGATCTTTCGAGGAGCACCACGCTACTCACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCACTGAGAGCTCCAGCGCTGGTGGCGCCACCAACGCTACTCACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTTCGCGAATGTTGTAAGTCTTCTCAGAGCTCTCAGAGTCTCAGTGGCGCCAC 3060
Db 3001 CCCTATGCTTTCGCGAATGTTGTAAGTCTTCTCAGAGCTCTCAGAGTCTCAGTGGCGCCAC 3060
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGC 3120
Db 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGT 3120
QY 3121 CGCAGCGAATACCTGTGCGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 GACTGCGCGGACCTTGTGCGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGCAAGTCTTTCTCAACTTCAGGCCATTTGGTCCGTCACCAACGTAACCTCAC 3240
Db 3181 CCAGAAATGTGCAAGTCTTTCTCAACTTCAGGCCATCTCGTCCGCGCACCACCAACGTAACCTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCGAGCGCGGCGAGTACCGGACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCGAGCGCGGCGGAGTACCGGACGAGTACCGGACTACGCT 3300

RESULT 6
US-10-257-384A-1
; Sequence 1, Application US/10257384A
; Publication No. US20040087524A1
; GENERAL INFORMATION:
; APPLICANT: Wiederanders, Bernd
; APPLICANT: Maubach, Gunter
; TITLE OF INVENTION: Agent for postoperative use after removal of bone tumors
; FILE REFERENCE: 2945-101
; CURRENT APPLICATION NUMBER: US/10/257,384A
; PRIOR FILING DATE: 2003-08-21
; PCT/DE 01/01510
; PRIOR FILING DATE: 2001-04-18
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 19
```

```
; SOFTWARE: PatentIn Ver. 2.1.1
; SEQ ID NO 1
; LENGTH: 7373
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Expression vector pMalc2chbmp2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)..(3453)
; OTHER INFORMATION: /product = fusion protein MBP Cystatin C BMP-2
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2689)
; OTHER INFORMATION: Faktor Xa Protease /position=388
; OTHER INFORMATION: (Amino acid sequence)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3067)..(3108)
; OTHER INFORMATION: /spacer between Cystatin C andBMP-2
; PUBLICATION INFORMATION:
; AUTHORS: Zwick, M B.
; TITLE: Expression vector pMal-X, complete sequence
; JOURNAL: Anal. Biochem.
; VOLUME: 264
; ISSUE: 1
; PAGES: 87-97
; DATE: 1998-11-01
; DATABASE ACCESSION NUMBER: Entrez Nucleotide database/AF031813
; DATABASE ENTRY DATE: 2001-05-07
; RELEVANT RESIDUES: 1528 - 3453
US-10-257-384A-1
```

```
Query Match 82.0%; Score 2705.8; DB 18; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAACTGATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAACTGATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGCGCCAGCGTTCCTCGGAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGCGCCAGCGTTCCTCGGAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTAATTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTAATTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGGTGCGACCTCAAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGGTGCGACCTCAAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTCGGGCGGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGGGCGGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
QY 361 TGTGTGTGTCGATGGTAGAACGAGCGGCTGGAAGCCCTGTAAAGCGGGGTGCACAAATC 420
Db 361 TGTGTGTGTCGATGGTAGAACGAGCGGCTGGAAGCCCTGTAAAGCGGGGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGGTCAGTGGGTGATCAATTAATCTCGGTGGATGACGAGATGCCA 480
Db 421 TTCTCGCGCAACCGGTCAGTGGGTGATCAATTAATCTCGGTGGATGACGAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACCTAAATGTTCCGCGGTTATTTCTTGATGTCTCTGACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACCTAAATGTTCCGCGGTTATTTCTTGATGTCTCTGACAGA 540
```


QY 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACCGGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACCGGACTGGCGTGGAGCATC 600
QY 601 TGGTCCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTCTGCG 660
Db 601 TGGTCCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTCTGCG 660
QY 661 CGCGTCTGGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 720
Db 661 CGCGTCTGGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 720
QY 721 CGGAACGGGAAGCGGACGTGGAGTGCATGTCCGGTTCCTCAACAAACCAATGCTGA 780
Db 721 CGGAACGGGAAGCGGACGTGGAGTGCATGTCCGGTTCCTCAACAAACCAATGCTGA 780
QY 781 ATGAGGGCATCTTCCCACTGCGATGCTGTGTCGAACGATCAGATGGCGCTGGCGCA 840
Db 781 ATGAGGGCATCTTCCCACTGCGATGCTGTGTCGAACGATCAGATGGCGCTGGCGCA 840
QY 841 TCGCGGCCATTACGAGTCCGGCTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
Db 841 TCGCGGCCATTACGAGTCCGGCTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
QY 901 ACGATACGGAACAGCTCATGTTATATCCCGCGTTAACCAACCAATCAACAGGATTTTC 960
Db 901 ACGATACGGAACAGCTCATGTTATATCCCGCGTTAACCAACCAATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCAGGSCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCAGGSCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGGACAGAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGGACAGAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGAGTGGAGCAACGAAATTAATGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGAGTGGAGCAACGAAATTAATGAGTTAGCTCACTCATTAG 1200
QY 1201 GCACAAATCTCATGTTTGAAGCTTATCATGCACTGCACGGTGCAACCAATGCTTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAAGCTTATCATGCACTGCACGGTGCAACCAATGCTTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGATGCTGTGCAAGTTCGTAATCATCTGCATTAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGATGCTGTGCAAGTTCGTAATCATCTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCGGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCGGACATCAACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAAAATATCTGAAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTAGCGGATAACAAATTTACAGAAACAGCCAGTCCGTTAGGTGTTTCAACA 1500
Db 1441 ATTGTAGCGGATAACAAATTTACAGAAACAGCCAGTCCGTTAGGTGTTTCAACA 1500
QY 1501 GCACCTCACCAACAGGACATAGATTATGAACCTGAAGAGGTAAACCTGATATCTGG 1560
Db 1501 GCACCTCACCAACAGGACATAGATTATGAACCTGAAGAGGTAAACCTGATATCTGG 1560
QY 1561 ATTTACCGGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTACCGGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTTAAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCACAGGTT 1680

Db 1621 ACCGGAATTTAAAGTCAACCGTTGAGCATCCGATAAATCTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGGCAACTGGCGATGGCCCTGACATATATCTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATATATCTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATAAAGAACTGAAAGGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAAGAACTGAAAGGAAAGTGAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACGCTACTTCAACCTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCAACCTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACACGCAAAAGTGAATTAATGGTTAAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGCAAAAGTGAATTAATGGTTAAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCAAACCATCAAAACCGTTGCTGGCTGCTGAGCGCAGGATTAACCGCCCGAGT 2340
Db 2281 AAGGTCAAACCATCAAAACCGTTGCTGGCTGCTGAGCGCAGGATTAACCGCCCGAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCAAAAGATCCACGTTATTCGCCCAACCAACGCGCAAGAAAGGTGAATCATG 2520
Db 2461 TTGGCAAAAGATCCACGTTATTCGCCCAACCAACGCGCAAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCGACGGTCTGTCAGCTGTCGATGAAGCGCTGAAGAGCGCGACACTAAATTCGAGCTCG 2640
Db 2581 GCGACGGTCTGTCAGCTGTCGATGAAGCGCTGAAGAGCGCGACACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAAACCAACCTTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAACAATAACAAACCAACCTTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
QY 2701 GGATCCTCTTCTGTTGGCCCGCGGCTTCGAGCGCGG 2741

Qy	1621	ACCGGAATTTAAAGTCA	CCGTTGAGCATCCGGATAA	CTGGAAGAGAAATTTCC	CACAGGTT	1680
Db	1621	ACCGGAATTTAAAGTCA	CCGTTGAGCATCCGGATAA	CTGGAAGAGAAATTTCC	CACAGGTT	1680
Qy	1681	GGGCAACTGGCGATG	CCCTGCACATTATCTT	CTGGGCACACGACCGCTTT	TGGTGGCTAC	1740
Db	1681	GGGCAACTGGCGATG	CCCTGCACATTATCTT	CTGGGCACACGACCGCTTT	TGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCTGT	TGGCTGAAATCACCC	CGSACAAACGTTTCC	AGGATGTAT	1800
Db	1741	GCTCAATCTGGCTGT	TGGCTGAAATCACCC	CGSACAAACGTTTCC	AGGATGTAT	1800
Qy	1801	CCGTTTACCTGGGATG	CGGTACGTTACAA	CGGCAAGCTGATTCG	TCTTACCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATG	CGGTACGTTACAA	CGGCAAGCTGATTCG	TCTTACCCGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGA	TTTATAACAAAGATCT	GTGTCGCCGAACCGCC	CAAAAAACCTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGA	TTTATAACAAAGATCT	GTGTCGCCGAACCGCC	CAAAAAACCTGGGAA	1920
Qy	1921	GAGATCCGGCGCTG	GTATAAAGACTGAA	AGCGAAAGGTAA	AGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGGCGCTG	GTATAAAGACTGAA	AGCGAAAGGTAA	AGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACGGTACT	TTCACCTGCGCCCTG	NTTCTGCTCACCGGGGTT	TATGCGTTCGTTCAAG	2040
Db	1981	CTGCAAGAACGGTACT	TTCACCTGCGCCCTG	NTTCTGCTCACCGGGGTT	TATGCGTTCGTTCAAG	2040
Qy	2041	TATGAAAAACGGCA	AGTAGTAAAGACG	TGGCGCTGGATAC	CGCTGGCGGGAAGCG	2100
Db	2041	TATGAAAAACGGCA	AGTAGTAAAGACG	TGGCGCTGGATAC	CGCTGGCGGGAAGCG	2100
Qy	2101	GGTCTGACTTCTG	GTGTGACCTGATTA	AAAAAACAACATG	ATCAGACACCGATTAC	2160
Db	2101	GGTCTGACTTCTG	GTGTGACCTGATTA	AAAAAACAACATG	ATCAGACACCGATTAC	2160
Qy	2161	TCCATCCGAGAGCT	CGCTTTTATAAAGG	CGGAACAGGATG	ACCATCAACGCCCGTGG	2220
Db	2161	TCCATCCGAGAGCT	CGCTTTTATAAAGG	CGGAACAGGATG	ACCATCAACGCCCGTGG	2220
Qy	2221	GCATGGTCCAACT	CGACACGAGCAAA	AGTGAATTTATGGT	TAAACGCTACTGCCGACCTTC	2280
Db	2221	GCATGGTCCAACT	CGACACGAGCAAA	AGTGAATTTATGGT	TAAACGCTACTGCCGACCTTC	2280
Qy	2281	AAGGGTCAACCAT	CAAAACCGTTT	CGTGGCGTGCTG	AGCGCAGGTAATTAACGCCCGCAGT	2340
Db	2281	AAGGGTCAACCAT	CAAAACCGTTT	CGTGGCGTGCTG	AGCGCAGGTAATTAACGCCCGCAGT	2340
Qy	2341	CCGAACAAAGAGCT	GGCAAAAGAGTCT	CTGAAAACCTATCT	GTGACTGATGAAGGTCG	2400
Db	2341	CCGAACAAAGAGCT	GGCAAAAGAGTCT	CTGAAAACCTATCT	GTGACTGATGAAGGTCG	2400
Qy	2401	GAAGCGGTTAATAA	AGACAAACCGCT	GGGTGCGGTAG	CGCTGAACTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAA	AGACAAACCGCT	GGGTGCGGTAG	CGCTGAACTCTTACGAGGAAGAG	2460
Qy	2461	TTGGGAAAAGATCA	CGATTTGGCGCC	ACACATGAAAAACG	CCCAAGAGGTGAATCATG	2520
Db	2461	TTGGGAAAAGATCA	CGATTTGGCGCC	ACACATGAAAAACG	CCCAAGAGGTGAATCATG	2520
Qy	2521	CCGAACATCCCG	CAGATGTCGCTTT	CTGTGGTATCCG	TGCTGCGGTATCAACGCC	2580
Db	2521	CCGAACATCCCG	CAGATGTCGCTTT	CTGTGGTATCCG	TGCTGCGGTATCAACGCC	2580
Qy	2581	GCCAGCGGTCTG	AGCTGTGATGAG	CCCTGAAAGACG	CGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTCTG	AGCTGTGATGAG	CCCTGAAAGACG	CGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAAACAAACA	ATAACATAACAA	CAACTCCGGAT	TCGAGGGAAGATTTTCAGAATTC	2700
Db	2641	AACAAACAAACA	ATAACATAACAA	CAACTCCGGAT	TCGAGGGAAGATTTTCAGAATTC	2700

Qy	2701	GGATCCTCT 2709		
Db	2701	GGATCCTCT 2709		
RESULT 8				
US-10-343-859-8				
; Sequence 8, Application US/10343859				
; Publication No. US20040110161A1				
; GENERAL INFORMATION:				
; APPLICANT: Nanogen Recognomics GMBH				
; TITLE OF INVENTION: Method for detecting mutations in				
; TITLE OF INVENTION: nucleotide sequences				
; FILE REFERENCE: 612,406-033				
; CURRENT APPLICATION NUMBER: US/10/343,859				
; CURRENT FILING DATE: 2003-11-24				
; PRIOR APPLICATION NUMBER: PCT/EP01/08127				
; PRIOR FILING DATE: 2001-07-13				
; PRIOR APPLICATION NUMBER: 10038237.1				
; PRIOR FILING DATE: 2000-08-04				
; NUMBER OF SEQ ID NOS: 52				
; SOFTWARE: Patent In Ver. 2.1				
; SEQ ID NO 8				
; LENGTH: 6648				
; TYPE: DNA				
; ORGANISM: Artificial sequence				
; FEATURE:				
; OTHER INFORMATION: Description of the artificial sequence: plasmid				
; OTHER INFORMATION: pMAL-c2x				
US-10-343-859-8				
Query Match 81.8%; Score 2701; DB 19; Length 6648;				
Best Local Similarity 99.8%; Pred. No. 0;				
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;				
Qy	1	CCGACACCATCGAATGTTGCAAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAAGAGA	60	
Db	1	CCGACACCATCGAATGTTGCAAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAAGAGA	60	
Qy	61	GTCAATTACAGGTGGTGAATGTGAAACCCAGTAACGTTTATACGATGTCGACAGATATGCCG	120	
Db	61	GTCAATTACAGGTGGTGAATGTGAAACCCAGTAACGTTTATACGATGTCGACAGATATGCCG	120	
Qy	121	GTGTCCTCTTATCAGACCGCTTTCCGCGTGGTGAAACGAGGCGAGCCACGTTTCTCGAAAA	180	
Db	121	GTGTCCTCTTATCAGACCGCTTTCCGCGTGGTGAAACGAGGCGAGCCACGTTTCTCGAAAA	180	
Qy	181	CGCGGGA AAAAGTGGAAAGCGCGCATGGCGAGCTGAATTCATTC CCAACCGCGTGGCAC	240	
Db	181	CGCGGGA AAAAGTGGAAAGCGCGCATGGCGAGCTGAATTCATTC CCAACCGCGTGGCAC	240	
Qy	241	AACAACATGGCGGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTGC	300	
Db	241	AACAACATGGCGGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTGC	300	
Qy	301	ACGGCCGCTCGCAAAATTGTCCGCGCGATTTAAATCTCGCGCCGATCAACTGGGTGCCAGCG	360	
Db	301	ACGGCCGCTCGCAAAATTGTCCGCGCGATTTAAATCTCGCGCCGATCAACTGGGTGCCAGCG	360	
Qy	361	TGTTGGTGTTCGATGGTTAGAACGAAGCGCGCTCGAAGCTGTGAAGCGCGGTGCACAATC	420	
Db	361	TGTTGGTGTTCGATGGTTAGAACGAAGCGCGCTCGAAGCTGTGAAGCGCGGTGCACAATC	420	
Qy	421	TTCTCGGCAACCGGTCAAGTGGGTGATCATTAATCTCGCTGGATGACCAAGATGCCA	480	
Db	421	TTCTCGGCAACCGGTCAAGTGGGTGATCATTAATCTCGCTGGATGACCAAGATGCCA	480	
Qy	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGCTTAATTTCTTGATGTTCTGACAGA	540	
Db	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGCTTAATTTCTTGATGTTCTGACAGA	540	
Qy	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACCGCATGGGGTGGAGCATC	600	

Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGTGGAGCATC 600
Qy TGGTCGCAATTGGGTACACAGCAAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db TGGTCGCAATTGGGTACACAGCAAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGCTCGG 660
Qy GCGGTCGCGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 720
Db GCGGTCGCGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 720
Qy CGGAACGGGAAGGACATGGAGTGCATGTCGGGTTTTCAACAAACCAATGCTGA 780
Db CGGAACGGGAAGGACATGGAGTGCATGTCGGGTTTTCAACAAACCAATGCTGA 780
Qy ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
Db ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
Qy TGC GCGCATTAACGAGTCCGCGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db TGC GCGCATTAACGAGTCCGCGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCACCATCAAAACAGATTTTC 960
Db ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCACCATCAAAACAGATTTTC 960
Qy GCCTGCTGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db GCCTGCTGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy AGGCAATCAGCTGTTGCCGTCTCATCTGGTGAAAGAAAAACCAACCTGCGGCCCAATA 1080
Db AGGCAATCAGCTGTTGCCGTCTCATCTGGTGAAAGAAAAACCAACCTGCGGCCCAATA 1080
Qy CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATCGAGTGGCAGCAGGTTT 1140
Db CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATCGAGTGGCAGCAGGTTT 1140
Qy CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCAGTCATTAG 1200
Db CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCAGTCATTAG 1200
Qy GCACAAATCTCATGTTGACAGCTTATCATCGATGCAAGTGCACCAATGCTTCTGGCG 1260
Db GCACAAATCTCATGTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Qy TCAGGACGCATCGAAGCTGTGGTATGCGTGCAGTGCATTAATCACTGCATAATTCG 1320
Db TCAGGACGCATCGAAGCTGTGGTATGCGTGCAGTGCATTAATCACTGCATAATTCG 1320
Qy TGTGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTTTCGCGCGACATCATACCGGTT 1380
Db TGTGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTTTCGCGCGACATCATACCGGTT 1380
Qy CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGGA 1440
Db CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGGA 1440
Qy ATTGTGAGCGGATAACAAATTTTCAACAGGAACAGCAGTCCGTTTTAGGTGTTTTTCAG 1500
Db ATTGTGAGCGGATAACAAATTTTCAACAGGAACAGCAGTCCGTTTTAGGTGTTTTTCAG 1500
Qy GCATTTCAACCAACGAGCATTAGATTATGAAATCTGAAGAGGTAACCTGGTAATCTGG 1560
Db GCATTTCAACCAACGAGCATTAGATTATGAAATCTGAAGAGGTAACCTGGTAATCTGG 1560
Qy ATTAACGGCGATAAGGCTATAACGGTCTCGCTGAACTCGGTAGAAATTCGAGAAAGAT 1620
Db ATTAACGGCGATAAGGCTATAACGGTCTCGCTGAACTCGGTAGAAATTCGAGAAAGAT 1620
Qy ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAACTGGAAGAGAAATTCGCCACAGGTT 1680
Db ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAACTGGAAGAGAAATTCGCCACAGGTT 1680

Qy 1681 GCGCAAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCCGTTTGGTGGCTAC 1740
Db 1681 GCGCAAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCCGTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGCAAAAGGTTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGCAAAAGGTTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAACGCTTATCGCTGATTATTAACAAAGATCTGCTGCGCAACCCGCAAAACCTGGGAA 1920
Db 1861 GAACGCTTATCGCTGATTATTAACAAAGATCTGCTGCGCAACCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAAACCGTACTTTCACCTGCGCGCTGATTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAAACCGTACTTTCACCTGCGCGCTGATTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAACAACACATGATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAACAACACATGATGACAGACCGGATTAC 2160
Qy 2161 TCCATCGCAAGAGTGCCTTTTAAATAAAGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCGCAAGAGTGCCTTTTAAATAAAGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Qy 2221 GCATGTGTCACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC 2280
Db 2221 GCATGTGTCACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC 2280
Qy 2281 AAGGTCACCAATCCAAACCGTTTCTGAGCGAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTTCTGAGCGAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTTCTCGAATACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTTCTCGAATACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGGAAAGATCCAGTATTTGCGGCAACATGGAAGAACCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGAAAGATCCAGTATTTGCGGCAACATGGAAGAACCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACAATCCGAGATGTCGGTCTTCTGTAATGCGGTGCTGCTGGGTGATCAACGCC 2580
Db 2521 CCGAACAATCCGAGATGTCGGTCTTCTGTAATGCGGTGCTGCTGGGTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGCTCAGACTGTGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTCAGACTGTGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGATTCAGAAATTC 2700
Qy 2701 GGATCTCT 2709
Db 2701 GGATCTCT 2709

RESULT 9

US-10-343-859-9
; Sequence 9, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; TITLE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9191
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-muts
; OTHER INFORMATION: Plasmid
US-10-343-859-9

Query Match 81.8%; Score 2698; DB 19; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGGGAAGAGA	60
DB	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGGGAAGAGA	60
QY	61	GTCAATTCAGGCTGGTGAATGTGAACAGTAACGTTATACCATGTCGACAGTATGCGG	120
DB	61	GTCAATTCAGGCTGGTGAATGTGAACAGTAACGTTATACCATGTCGACAGTATGCGG	120
QY	121	GTGCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACAGGCCAGCCACGTTTCTGGGAAAA	180
DB	121	GTGCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACAGGCCAGCCACGTTTCTGGGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAATCCAAACCGCGTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAATCCAAACCGCGTGGCAC	240
QY	241	AACAACCTGGCGGCAACAGTCTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACCTGGCGGCAACAGTCTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG	360
DB	301	ACGCGCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG	360
QY	361	TGGTGGTGTGATGTAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC	420
DB	361	TGGTGGTGTGATGTAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTATCCGTGGATGACAGGATGCCA	480
DB	421	TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTATCCGTGGATGACAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA	540
DB	481	TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACCGGTACGCGACTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACCGGTACGCGACTGGCGGTGGAGCATC	600
QY	601	TGGTCCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCATTAAGTTCTGTCTCGG	660
DB			

DB	601	TGGTCCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660
QY	661	CGCTCTGCGTCTGCGTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG	720
DB	661	CGCTCTGCGTCTGCGTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG	720
QY	721	CGAAACGGGAAGCGCACTGGAGTGCCTATGTCGGGTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGAAACGGGAAGCGCACTGGAGTGCCTATGTCGGGTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGCGCATCGTTCCCACTGCGATGCTGTTGTCACACGATCAGATGCGCTGGGCGCAA	840
DB	781	ATGAGGCGCATCGTTCCCACTGCGATGCTGTTGTCACACGATCAGATGCGCTGGGCGCAA	840
QY	841	TGCGCGCATTTACCGAGTCCGGGCTGGCGGTGGTGGCGATATCTCGTAGTGGGATACG	900
DB	841	TGCGCGCATTTACCGAGTCCGGGCTGGCGGTGGTGGCGATATCTCGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTAAACACCATCAAAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTAAACACCATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGTGA	1020
DB	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCCCTCTCATCTGTTGAAAGAAAAACCAACCTCGGCGCCAAATA	1080
DB	1021	AGGCAATCAGCTGTTGCCCTCTCATCTGTTGAAAGAAAAACCAACCTCGGCGCCAAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATCGAGTGGCAGCAGCAGGTTT	1140
DB	1081	CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATCGAGTGGCAGCAGCAGGTTT	1140
QY	1141	CCGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
DB	1141	CCGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
QY	1201	GCACAATTTCTCATGTTGACAGCTTATCATCGATGCGTGCACCAATGTTCTTGGCG	1260
DB	1201	GCACAATTTCTCATGTTGACAGCTTATCATCGATGCGTGCACCAATGTTCTTGGCG	1260
QY	1261	TCAGCAGCCATCGGAAGCTGTGTTGCTGATGCTGTCAGCTGTAATCACTCATTAATTCG	1320
DB	1261	TCAGCAGCCATCGGAAGCTGTGTTGCTGATGCTGTCAGCTGTAATCACTCATTAATTCG	1320
QY	1321	TGTCCTCAAGCGCAGCTCCCGTTCTGGAATAATGTTTTTGGCGCGACATCATACCGTT	1380
DB	1321	TGTCCTCAAGCGCAGCTCCCGTTCTGGAATAATGTTTTTGGCGCGACATCATACCGTT	1380
QY	1381	CTGGCAAAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
DB	1381	CTGGCAAAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
QY	1441	ATTGTAGCGGATTAACAAATTTTCAACAGAAACAGCCAGTCCGTTAGGTTGTTTTCACGA	1500
DB	1441	ATTGTAGCGGATTAACAAATTTTCAACAGAAACAGCCAGTCCGTTAGGTTGTTTTCACGA	1500
QY	1501	GCACCTTCCCAACAGGACCATAGATTATGAAAATGAAAATGAAAATGAAAATGAAAATG	1560
DB	1501	GCACCTTCCCAACAGGACCATAGATTATGAAAATGAAAATGAAAATGAAAATGAAAATG	1560
QY	1561	ATTAACGGCGATTAAGGCTATAACCGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAACGGCGATTAAGGCTATAACCGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCACAGGTT	1680
DB	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCACAGGTT	1680
QY	1681	GGGCAATCGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTCTGCTGCTAC	1740
DB	1681	GGGCAATCGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTCTGCTGCTAC	1740

Qy	1741	GCTCAATCTGGCCTGTGGCTGAAATCACCCCGACAAGCGTTCCAGGACAAGCTGAT	1800
Db	1741	GCTCAATCTGGCCTGTGGCTGAAATCACCCCGACAAGCGTTCCAGGACAAGCTGAT	1800
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTACAA CGGCAAGCTGATTGCTTACC GATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCGGTACGTTACAA CGGCAAGCTGATTGCTTACC GATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCCGCCAAAACCTCGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCCGCCAAAACCTCGGAA	1920
Qy	1921	GAGATCCCGGGCTGGATAAGAACTGAAGCGAAGAGGTGAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGGGCTGGATAAGAACTGAAGAGCGAAGAGGTGAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTTCACCTGSCCGCTGATTGCTGCTGACGGGGGTATGCGTTCAG	2040
Db	1981	CTGCAAGAACCGTACTTTCACCTGSCCGCTGATTGCTGCTGACGGGGGTATGCGTTCAG	2040
Qy	2041	TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATAACGCTGCGCGGAAGCG	2100
Db	2041	TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATAACGCTGCGCGGAAGCG	2100
Qy	2101	GGTCTGACCTTCTGGTTGACCTGATTAAAAACAACACATGATGAATCGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGGTTGACCTGATTAAAAACAACACATGATGAATCGAGACACCGATTAC	2160
Qy	2161	TCCATCGCAAGAGCTGCCTTTAATAAAGCGAAACAGCGATGACCATCAACGCCCGCTGG	2220
Db	2161	TCCATCGCAGAGCTGCCTTTAATAAAGCGAAACAGCGATGACCATCAACGCCCGCTGG	2220
Qy	2221	GCATGFTCCAACTCGACACCGACAAAGTGAATTTATGGTTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGFTCCAACTCGACACCGACAAAGTGAATTTATGGTTAAACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGGTCAACCATCCCAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTACGCCGCGCAGT	2340
Db	2281	AAGGGTCAACCATCCCAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTACGCCGCGCAGT	2340
Qy	2341	CCGAAACAAGAGCTGGCAAAAGAGTTCTCCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAAACAAGAGCTGGCAAAAGAGTTCTCCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
Qy	2401	GAGCGGTTAATAAGACAAACCGTGGTGGCTGAGCGCTGAAAGTCTTACGAGGAAGAG	2460
Db	2401	GAGCGGTTAATAAGACAAACCGTGGTGGCTGAGCGCTGAAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAAGATCCACGTATTTCGCCGCCATCGAAACCGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTATTTCGCCGCCATCTATGGAAAAACGCCAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGCTTTCTGTTATCGCTGCGTACTCGCGTGTATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGCTTTCTGTTATCGCTGCGTACTCGCGTGTATCAACGCC	2580
Qy	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAACAACAATTAACAATAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2641	AACAACAACAATTAACAATAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Qy	2701	GGATCC 2706	
Db	2701	GGATCC 2706	

RESULT 10

US-10-263-153-40

Sequence 40, Application US/10263153

```

; Publication No. US20040067239A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
;
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN.
;
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
;
; FILE REFERENCE: 6984.US.O1
;
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
;
; NUMBER OF SEQ ID NOS: 74
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 40
;
; LENGTH: 7112
;
; TYPE: DNA
;
; ORGANISM: Toxoplasma gondii
;
; FEATURE:
;
; NAME/KEY: CDS
;
; LOCATION: (1528)...(3189)
;
; OTHER INFORMATION: pMBP-c2X-Toxop30del11 (52-214aa)
;
; US-10-263-153-40

```

Query Match	81.5%	Score 2690.4	DB 18	Length 7112
Best Local Similarity	99.8%	Pred. No. 0		
Matches 2694	Conservative 0	Mismatches 6	Indels 0	Gaps 0

QY	1	CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA	60
DB	1		
QY	61	GTCAAATTCAGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTGCGAGAGTATGCCG	120
DB	61		
QY	121	GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGACGACCAAGTTCTTCGCGAAA	180
DB	121	GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGACGACCAAGTTCTTCGCGAAA	180
QY	181	CGCGGAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCGAACCGCGTGGCAC	240
DB	181	CGCGGAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCGAACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCGCAAAACAGTCTGCTGAATGGCGGTGGCACCTCCAGTCTCGCCCTGC	300
DB	241	AACAACTGGCGGCGCAAAACAGTCTGCTGAATGGCGGTGGCACCTCCAGTCTCGCCCTGC	300
QY	301	ACGCGCGTGCAGAAATGTTCGGCGGATTAATPCTCGCGCGATCAACTGGGTGCCACGC	360
DB	301	ACGCGCGTGCAGAAATGTTCGGCGGATTAATPCTCGCGCGATCAACTGGGTGCCACGC	360
QY	361	TGGTGTGTTCGATGTTAGAACGAGCGGCTCGAAGCTGTAAAGCGCGGTGCACAATC	420
DB	361	TGGTGTGTTCGATGTTAGAACGAGCGGCTCGAAGCTGTAAAGCGCGGTGCACAATC	420
QY	421	TTCTCCGCGAAACCGTFCAGTGGCTCATCATTAACCTATCGCTGGATGACACGAGTGCCA	480
DB	421	TTCTCCGCGAAACCGTFCAGTGGCTCATCATTAACCTATCGCTGGATGACACGAGTGCCA	480
QY	481	TTGCTGTGGAAGTGCTGCACCTAAATGTTCCGCGGTTATTTCTTGATGTCTCTGACAGA	540
DB	481	TTGCTGTGGAAGTGCTGCACCTAAATGTTCCGCGGTTATTTCTTGATGTCTCTGACAGA	540
QY	541	CACCCATCAACAGTATATTTTCTCCCATGAAGA CGGTACGCACTGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATATTTTCTCCCATGAAGA CGGTACGCACTGGGCGTGGAGCATC	600
QY	601	TGTCGCAATGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG	660
DB	601	TGTCGCAATGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCGATAG	720

Db 661 |||||CGGCTCTCGGCTGGCGTAAATATCTCCTCGCAATCAAAATCAGCGGATAG 720
Qy 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 |||||CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGATCTGTTCCCATCTGGATGCTGGTTGCCAACAGATCAGATGCGGTGGCGCGAA 840
Db 781 ATGAGGGATCTGTTCCCATCTGGATGCTGGTTGCCAACAGATCAGATGCGGTGGCGCGAA 840
Qy 841 TGGCGGCATTACCGAGTCCGGGCTCGCGTTGGTGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCATTACCGAGTCCGGGCTCGCGTTGGTGCGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGTGGACCGTCTGCTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGTCTGCTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCGCTCTCAGTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCAGTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGGCAGCACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGGCAGCACAGGTTT 1140
Qy 1141 CCCGACTGGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGTACTCACTCATTTAG 1200
Db 1141 CCCGACTGGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGTACTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTCAGACGTTATCATGACGTGCACGCGTGCAACCAATGCTTCGGG 1260
Db 1201 GCACAAATCTCATGTTTCAGACGTTATCATGACGTGCACGCGTGCAACCAATGCTTCGGG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCTGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCTGTAATCACTGCATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCATCTCCGTTCTCGATTAATGTTTTTGGCGCACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCATCTCCGTTCTCGATTAATGTTTTTGGCGCACATCATACGGTT 1380
Qy 1381 CTGGCAATATCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAATATCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATACAAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTCAGGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTCAGGA 1500
Qy 1501 GCATCTCCACAAAGGACCATAGCATATGAAATCGAAGAGGTAAACTGCTAATCTGG 1560
Db 1501 GCATCTCCACAAAGGACCATAGCATATGAAATCGAAGAGGTAAACTGCTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 |||||GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAAGCTGAT 1800

Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAAGCTGAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACCGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACCGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Qy 1861 GAACGGTATCGCTGATTTATTAACAAAGATCTGCTGCGGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAACGGTATCGCTGATTTATTAACAAAGATCTGCTGCGGAACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGAATAAGAACTGAAAGCGAAAGTGAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGAATAAGAACTGAAAGCGAAAGTGAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATCAAAACCGCAAGTACGATTAAGAGACGTGGCGGTGGATAACGCTGGCGGCAAGCG 2100
Db 2041 TATCAAAACCGCAAGTACGATTAAGAGACGTGGCGGTGGATAACGCTGGCGGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGGTG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGGTG 2220
Qy 2221 GCATGTTCCAAATCGACACAGCAAGTGAATTAATGTTGTTAAACGTTACTGCGACCTTC 2280
Db 2221 GCATGTTCCAAATCGACACAGCAAGTGAATTAATGTTGTTAAACGTTACTGCGACCTTC 2280
Qy 2281 AAGGTCAACCATCAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACCGCCCACT 2340
Db 2281 AAGGTCAACCATCAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACCGCCCACT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAACGGTATAAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAACGGTATAAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATTCGCCCAACCATGGAACACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTCGCCCAACCATGGAACACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCGAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCGAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Qy 2641 AACCAACCAACAAATAAATAAACAACCTCGGATCGAGGGAAGGATTTTCAAAATTC 2700
Db 2641 AACCAACCAACAAATAAATAAACAACCTCGGATCGAGGGAAGGATTTTCAAAATTC 2700

RESULT 11
US-10-263-153-30
; Sequence 30, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.

; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 7259
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3336)
; OTHER INFORMATION: pMBP-c2X-Toxop30del14del18 (83-294aa)
US-10-263-153-30

Query Match 81.5%; Score 2690.4; DB 18; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGA 60

Qy 61 GTCAATTGAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTGAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATGCCG 120

Qy 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCGCAGCAGTTCGCGAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCGCAGCAGTTCGCGAAAA 180

Qy 181 CGCGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAATTAACATCCCAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGGCAACAGTCGTTGTGATGTGGCTTGCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAACAGTCGTTGTGATGTGGCTTGCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGCGCGTGCAGAAATGTGCGCGGATTAATCTCGCGCGGATCACTGGGTGCGACG 360
Db 301 ACGCGCGTGCAGAAATGTGCGCGGATTAATCTCGCGCGGATCACTGGGTGCGACG 360

Qy 361 TGGTGGTGCAGTGTAGAACGAGCGCGTGAAGCGCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGCAGTGTAGAACGAGCGCGTGAAGCGCTGTAAAGCGCGGTGCACAATC 420

Qy 421 TTCTCGCGCAACCGGTGAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGGTGAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480

Qy 481 TTGCTGTGGAAGCTGCGTGCACATAATGTCGCGGTATTTCTTGATGTCCTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCGTGCACATAATGTCGCGGTATTTCTTGATGTCCTGACACAGA 540

Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGAGCGGTACGCGACTGGGCGTGGAGCATC 600

Qy 601 TGGTGCATTTGGGTACACAGCAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660
Db 601 TGGTGCATTTGGGTACACAGCAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660

Qy 661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCCCAATCAAAATCAGCCCATAG 720
Db 661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCCCAATCAAAATCAGCCCATAG 720

Qy 721 CGGAAACGGGAAGCGGACTGAGTGCCATGTCGGTTCGAAACCAACATGCAATGCTGA 780
Db 721 CGGAAACGGGAAGCGGACTGAGTGCCATGTCGGTTCGAAACCAACATGCAATGCTGA 780

Qy 781 ATGAGGCGATCGTTCCCACTGCGATGCTGGTTCGAACGATCAGATCGCGCTGGCGCAA 840
Db 781 ATGAGGCGATCGTTCCCACTGCGATGCTGGTTCGAACGATCAGATCGCGCTGGCGCAA 840

Qy 841 TGGCGCCATTACCGAGTCCGGGCTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTGGTGGGATATCTCGGTAGTGGGATACG 900

Qy 901 ACGNATCCGAAGACAGCTCATGTTATATCCGCGGTTAACACCATCAACAGATTTTC 960
Db 901 ACGNATCCGAAGACAGCTCATGTTATATCCGCGGTTAACACCATCAACAGATTTTC 960

Qy 961 GCCTGTGGGGCAAAACAGCGTGGACCGTTCGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACAGCGTGGACCGTTCGCAACTCTCTCAGGGCCAGCGGTGA 1020

Qy 1021 AGGCAATACGTGTTGCCGCTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATACGTGTTGCCGCTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATGAGCTGGGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATGAGCTGGGACGACAGGTTT 1140

Qy 1141 CCGGACTGGAAAGCGGCGAGTGAAGCAACGCAATTAATGAGTTCAGTCTCATTTAG 1200
Db 1141 CCGGACTGGAAAGCGGCGAGTGAAGCAACGCAATTAATGAGTTCAGTCTCATTTAG 1200

Qy 1201 GCAAAATTCATGTTTGAAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCAAAATTCATGTTTGAAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGCG 1260

Qy 1261 TCAGGCGACCATCGGAAGCTGCTGATGGTGGTGGAGTCTGTAATCAGTCAATTCG 1320
Db 1261 TCAGGCGACCATCGGAAGCTGCTGATGGTGGTGGAGTCTGTAATCAGTCAATTCG 1320

Qy 1321 TGTGCTCAAGCGCGACTCCCGTTCGGAATAATGTTTTTTCGCGCGACATCAACCGTT 1380
Db 1321 TGTGCTCAAGCGCGACTCCCGTTCGGAATAATGTTTTTTCGCGCGACATCAACCGTT 1380

Qy 1381 CTGGCAAAATTCGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGA 1440
Db 1381 CTGGCAAAATTCGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGA 1440

Qy 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA 1500

Qy 1501 GCACTTCAACCAACAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACTTCAACCAACAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAATCTGG 1560

Qy 1561 ATTAAACGGGATTAAGGCTTAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATTAAGGCTTAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTCGGAATTAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAAGAT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTCGGAATTAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAAGAT 1680

Qy 1681 GGGCAACTGGGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740

Qy 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCGGCAAAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCGGCAAAAGCGTTCAGGACAAAGCTGTAT 1800

Qy 1801 CCGTTTACCTGGGATGCGCGTACGTTTACAAACGCGCAAGCTGATTCCTCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGCGTACGTTTACAAACGCGCAAGCTGATTCCTCCGATCGCTGT 1860

Qy 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGTGCGCGAAACCGCGCAAAACCTGGGAA 1920

```
Db 1861 GNAAGGTTATCGCTGATTTATTAACAAGATCTGTCGGAACCCGCAAAAACCTTGGGA 1920
Qy 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGAGCGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGAGCGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGTGAATAACGCTGGCGCAAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGTGAATAACGCTGGCGCAAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACACATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACACATGAATGACACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCCCTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGTTGAACGCTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGTTGAACGCTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACACCATCAACCGTTCTGCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCATCAACCGTTCTGCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Qy 2341 CGAACAAGAGCTGGCAAAAGAGTTCCTGGAACACTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CGAACAAGAGCTGGCAAAAGAGTTCCTGGAACACTATCTGCTGACTGATGAAGGCTG 2400
Qy 2401 GAAAGCGTTAATAAAGACAAACCGCTGGTGGCGTGAAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAAGCGTTAATAAAGACAAACCGCTGGTGGCGTGAAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCACGATTTGCGCGCCACATGGAAGAGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGCGCCACATGGAAGAGCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAATATCCCGCAGATGTCGCTTTCTGTTGATGCGGTGATGCGGTGATCAACGCC 2580
Db 2521 CCGAATATCCCGCAGATGTCGCTTTCTGTTGATGCGGTGATGCGGTGATCAACGCC 2580
Qy 2581 GCCAGGCTGTCGACACTGTCGATGAAGCCCTGGAAGAGCGCGCAGACTAATTCGAGCTG 2640
Db 2581 GCCAGGCTGTCGACACTGTCGATGAAGCCCTGGAAGAGCGCGCAGACTAATTCGAGCTG 2640
Qy 2641 AACACAACAAATAAACAATAACAATAACAACCTCGGGATCGAGGGAAGNTTCAAGATTC 2700
Db 2641 AACACAACAAATAAACAATAACAATAACAACCTCGGGATCGAGGGAAGNTTCAAGATTC 2700
```

RESULT 12

```
US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 35
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528) ... (3399)
; OTHER INFORMATION: pMBP-c2X-Toxop30del10 (52-284aa)
US-10-263-153-35
```

Query Match

```
Best Local Similarity 81.5%; Score 2690.4; DB 18; Length 7322;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 CCGACACCATCGAATGGTGTGCAAAAACCTTTCCGGGTATGCGCATGATAGCCCGCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGTGCAAAAACCTTTCCGGGTATGCGCATGATAGCCCGCGGAAGAGA 60
Qy 61 GTCAATTCCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCGG 120
Db 61 GTCAATTCCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCGG 120
Qy 121 GTGTCCTTATCAGACCCGTTTCCCGCGTGTGAACCCAGGCCAGCCACGTTTCTGCGAAAA 180
Db 121 GTGTCCTTATCAGACCCGTTTCCCGCGTGTGAACCCAGGCCAGCCACGTTTCTGCGAAAA 180
Qy 181 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAACAGTCGTTGCTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAACAGTCGTTGCTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGCTCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCGCAGCG 360
Db 301 ACGCGCGCTCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCGCAGCG 360
Qy 361 TGGTGGTGTGATGGAACGAAGCGCGCTGAAGCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTGTGATGGAACGAAGCGCGCTGAAGCTGTAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACCGGTGAGTGGGTGATCAATTAATTAATTAATTAATTAATTAATTAATTA 480
Db 421 TTCTCGCGCAACCGGTGAGTGGGTGATCAATTAATTAATTAATTAATTAATTAATTAATTA 480
Qy 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGAGCGGTACGCGACTGGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGAGCGGTACGCGACTGGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCGTGTCGG 660
Db 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCGTGTCGG 660
Qy 661 CGCGTCTGCGCTGCGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGCTGCGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGGGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGGGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGCTTCCCACTCGCATGCTGGTTGCCAAGCATGATGCGGCTGGCGCGCA 840
Db 781 ATGAGGGCATGCTTCCCACTCGCATGCTGGTTGCCAAGCATGATGCGGCTGGCGCGCA 840
Qy 841 TGGCGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTGTGAGGATAG 900
Db 841 TGGCGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTGTGAGGATAG 900
```

Qy	901	ACGATACCGAAGACAGCTCATGTTTATATATCCCGCGTTTAAACACCATCAAAACAGGATTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACACCATCAAAACAGGATTTC	960
Qy	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTGTCGCCGCTCACTGCTGGAAGAAAGAAACACCACTTGGCGGCCAATA	1080
Db	1021	AGGGCAATCAGCTGTGTCGCCGCTCACTGCTGGAAGAAAGAAACACCACTTGGCGGCCAATA	1080
Qy	1081	CGCAAAACCGCTCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGCGTTT	1140
Db	1081	CGCAAAACCGCTCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGCGTTT	1140
Qy	1141	CCCGACTGGAAAGCGGGCAGTGACGGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Db	1141	CCCGACTGGAAAGCGGGCAGTGACGGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Qy	1201	GCACAAATCTCATGTTTGTGACAGCTTATCATCGACTGCACCGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGTGACAGCTTATCATCGACTGCACCGTGCACCAATGCTTCTGGCG	1260
Qy	1261	TCAGGAGCCATCGGAAGCTGTGTTATGCTGTGCAGGTGTAATCACTGCATAATTCG	1320
Db	1261	TCAGGAGCCATCGGAAGCTGTGTTATGCTGTGCAGGTGTAATCACTGCATAATTCG	1320
Qy	1321	TGTCGCTCAAGGCGCACTCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCAACCGGTT	1380
Db	1321	TGTCGCTCAAGGCGCACTCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCAACCGGTT	1380
Qy	1381	CTGGCAATAATCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTAATATGTGTGA	1440
Db	1381	CTGGCAATAATCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTAATATGTGTGA	1440
Qy	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCAACGA	1500
Db	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCAACGA	1500
Qy	1501	GCATTTACCAACAAAGGACCATAGATTATGAAATCTGAAAGGTAACTGGTAATCTCG	1560
Db	1501	GCATTTACCAACAAAGGACCATAGCATATGAAATCTGAAAGGTAACTGGTAATCTCG	1560
Qy	1561	ATTAAACGGGATTAAGGCTATAACGGTCTCGCTGGAAGTGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGGATTAAGGCTATAACGGTCTCGCTGGAAGTGGTAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGATAAGAAATTCGAGAAAGAT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGATAAGAAATTCGAGAAAGAT	1680
Qy	1681	GGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCGACACGACCGCTTTGGTGCTAC	1740
Db	1681	GGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCGACACGACCGCTTTGGTGCTAC	1740
Qy	1741	GCTCAATCTGGCCGTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCGTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCCGTACGTTTAAACGGCAAGCTGATTGCTTACCCCGATCCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCCGTACGTTTAAACGGCAAGCTGATTGCTTACCCCGATCCGCTGTT	1860
Qy	1861	GAAGCGTTATCGTGATTTATAACAAAGATCTGCTGCCGAAACCCGCGCAAAACCTGGGAA	1920
Db	1861	GAAGCGTTATCGTGATTTATAACAAAGATCTGCTGCCGAAACCCGCGCAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980

Qy	1981	CTGCAAGAACCGTATCTTCA	CTCTGCGCGCTGANTGCTGCT	GACGGGGGTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTATCTTCA	CTCTGCGCGCTGANTGCTGCT	GACGGGGGTATGCGTTCAAG	2040
Qy	2041	TATGAAAAAGCGCAAGTACGACAT	TAAAGACGTGGCGCTGGATACGCT	GCGGCGGAAGCG	2100
Db	2041	TATGAAAAAGCGCAAGTACGACAT	TAAAGACGTGGCGCTGGATACGCT	GCGGCGGAAGCG	2100
Qy	2101	GGTCTGACCTTCTCTGGTGTGAC	CTTGATTAATAAAACAAACACAT	GAATGACAGACACCGGATTAC	2160
Db	2101	GGTCTGACCTTCTCTGGTGTGAC	CTTGATTAATAAAACAAACACAT	GAATGACAGACACCGGATTAC	2160
Qy	2161	TCCATTCGACGAAGCTGCTTTTAA	TAAAGGCGAAACAGCGATGACCAT	CAACAGCGCCCGTGG	2220
Db	2161	TCCATTCGACGAAGCTGCTTTTAA	TAAAGGCGAAACAGCGATGACCAT	CAACAGCGCCCGTGG	2220
Qy	2221	GCATGGTCCAACTCGACACACAG	CAACAAAGTGAATTTATGGTTAA	CCGGTACTTGCACGACCTTC	2280
Db	2221	GCATGGTCCAACTCGACACACAG	CAACAAAGTGAATTTATGGTTAA	CCGGTACTTGCACGACCTTC	2280
Qy	2281	AAGGGTCAACCATCCRAACCGTT	CTGTGGCGTCTGAGCGCAGGTAT	TAAACGCGCGCAGT	2340
Db	2281	AAGGGTCAACCATCCRAACCGTT	CTGTGGCGTCTGAGCGCAGGTAT	TAAACGCGCGCAGT	2340
Qy	2341	CCGAAACAAAGAGCTGGCAAAAG	AGTTCTCTCGAAAACTATCTGCT	GACTGATGAAGGTCG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAAG	AGTTCTCTCGAAAACTATCTGCT	GACTGATGAAGGTCG	2400
Qy	2401	GAACGGGTTAATAAGACAAACCG	CTGGGTGCGGTAGCGCTGAAGCT	TTCACGAGGAAG	2460
Db	2401	GAACGGGTTAATAAGACAAACCG	CTGGGTGCGGTAGCGCTGAAGCT	TTCACGAGGAAG	2460
Qy	2461	TTGCGAAAGATCCACGTATTGCG	CGCACCATGGAATAACGCCAGAA	AGGTGAATCATG	2520
Db	2461	TTGCGAAAGATCCACGTATTGCG	CGCACCATGGAATAACGCCAGAA	AGGTGAATCATG	2520
Qy	2521	CCGAACTCCCGCAGATGTCCGGT	TTCTGGTATGCCGTGGCTACTC	GCGTGAATCAACGCC	2580
Db	2521	CCGAACTCCCGCAGATGTCCGGT	TTCTGGTATGCCGTGGCTACTC	GCGTGAATCAACGCC	2580
Qy	2581	GCCAGCGGTGCTCAGACTGTGAT	GAAGCCCTGAAAGACGCCGAG	ACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGCTCAGACTGTGAT	GAAGCCCTGAAAGACGCCGAG	ACTAATTCGAGCTCG	2640
Qy	2641	AACAAACAACTAATTAACAAAC	AACTTCGGGATCGAGGGAAG	ATTCAGAAATTC	2700
Db	2641	AACAAACAACTAATTAACAAAC	AACTTCGGGATCGAGGGAAG	ATTCAGAAATTC	2700

RESULT 13
US-10-263-153-25
; Sequence 25, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 7352
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528) ... (3429)

; OTHER INFORMATION: pMBP-c2X-ToxoP3odel14C (52-294aa) US-10-263-153-25									
Query Match 81.5%; Score 2690.4; DB 18; Length 7352; Best Local Similarity 99.8%; Pred. No. 0; Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;									
Qy	1	CGGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGCGCATGATGAGCCCGGAAGAGA	60					
Db	1	CGGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGCGCATGATGAGCCCGGAAGAGA	60					
Qy	61	GTCAATTCAGGGTGTGAATGTGA	ACCAGTAACGTATACGATCTCCAGAGTATGCCG	120					
Db	61	GTCAATTCAGGGTGTGAATGTGA	ACCAGTAACGTATACGATCTCCAGAGTATGCCG	120					
Qy	121	GTGTCCTTATCAGACCGGTTTC	CCGCGTGGTGAACCGACCGCACGCTTTCTCGGAAAA	180					
Db	121	GTGTCCTTATCAGACCGGTTTC	CCGCGTGGTGAACCGACCGCACGCTTTCTCGGAAAA	180					
Qy	181	CGCGGAAAAAGTGAAGCGGGAT	GGCGAGCTGAATTAATCCCAACCGCGTGGCAC	240					
Db	181	CGCGGAAAAAGTGAAGCGGGAT	GGCGAGCTGAATTAATCCCAACCGCGTGGCAC	240					
Qy	241	AACAACTGCGGGCAACAGTCTG	TTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300					
Db	241	AACAACTGCGGGCAACAGTCTG	TTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300					
Qy	301	ACGCGCCGTCGCAAAATTTG	TCGGCGGATTAATCTCGGCGCGATCAACTGGGTGCCAGCG	360					
Db	301	ACGCGCCGTCGCAAAATTTG	TCGGCGGATTAATCTCGGCGCGATCAACTGGGTGCCAGCG	360					
Qy	361	TGGTGGTTCGATGTGAGAAC	GAAGCGGGTCCGAAGCCTGTAAAGCGGCGTGCAATC	420					
Db	361	TGGTGGTTCGATGTGAGAAC	GAAGCGGGTCCGAAGCCTGTAAAGCGGCGTGCAATC	420					
Qy	421	TTCTCGCGCAACGGCTCAGT	GGCTGATCAATTAATCTCCGTGGATGACCAAGATGCCA	480					
Db	421	TTCTCGCGCAACGGCTCAGT	GGCTGATCAATTAATCTCCGTGGATGACCAAGATGCCA	480					
Qy	481	TTGCTGTGGAAAGTCCTGCA	TAAATGTTCCGGCGTTATTTCTGATGTTCTGACCCAGA	540					
Db	481	TTGCTGTGGAAAGTCCTGCA	TAAATGTTCCGGCGTTATTTCTGATGTTCTGACCCAGA	540					
Qy	541	CACCCATCAACAGTATATTT	CTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600					
Db	541	CACCCATCAACAGTATATTT	CTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600					
Qy	601	TGGTCGCATTTGGTACCAG	CAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660					
Db	601	TGGTCGCATTTGGTACCAG	CAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660					
Qy	661	CGCGTCTGCTGCTGGCTG	GGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720					
Db	661	CGCGTCTGCTGCTGGCTG	GGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720					
Qy	721	CGGAAACGGGAAGCGAC	TGGAGTGCATGTCGGTTTTTCAACAAACCAATGCTGA	780					
Db	721	CGGAAACGGGAAGCGAC	TGGAGTGCATGTCGGTTTTTCAACAAACCAATGCTGA	780					
Qy	781	ATGAGGGCATCGTTCCCA	CTGCGATGCTGGTTGCCAACGATCGAGTGGCGTGGCGGCA	840					
Db	781	ATGAGGGCATCGTTCCCA	CTGCGATGCTGGTTGCCAACGATCGAGTGGCGTGGCGGCA	840					
Qy	841	TGCGGCCATTAACGAGT	CCGGCTCGCGTGGTGGGATATCTCGTGGGATAGC	900					
Db	841	TGCGGCCATTAACGAGT	CCGGCTCGCGTGGTGGGATATCTCGTGGGATAGC	900					
Qy	901	ACGATACCGAAGACAG	CTCATGTATATCCCGCGTTAACACCAATCAAAACAGGATTTTC	960					
Db	901	ACGATACCGAAGACAG	CTCATGTATATCCCGCGTTAACACCAATCAAAACAGGATTTTC	960					
Qy	961	GCCTGCTGGGGCAAC	CCAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020					
Db	2041	TATGAAAAACGGCAAGT	ACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGCGAAAGCG	2100					
Db	2041	TATGAAAAACGGCAAGT	ACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGCGAAAGCG	2100					

Db	961	GCCTGCTGGGGCAAC	CAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020					
Qy	1021	AGGGCAATCAGCTGTG	TGCCCGTCTCAGCTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080					
Db	1021	AGGGCAATCAGCTGTG	TGCCCGTCTCAGCTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080					
Qy	1081	CGAAAAACCGCTCT	CCCCCGCGTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140					
Db	1081	CGAAAAACCGCTCT	CCCCCGCGTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140					
Qy	1141	CCGACTGCGAAAGCG	GGGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG	1200					
Db	1141	CCGACTGCGAAAGCG	GGGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG	1200					
Qy	1201	GCACAAATTTCTCAT	GTGTCAGAGCTTATCATCGATGCGTGCACCAATGCTTCTGCGG	1260					
Db	1201	GCACAAATTTCTCAT	GTGTCAGAGCTTATCATCGATGCGTGCACCAATGCTTCTGCGG	1260					
Qy	1261	TCAGGCAGCCATCG	GAAGCTGTGTATGGCTGTGCGAGTCTGTAATCACTGCATATTCG	1320					
Db	1261	TCAGGCAGCCATCG	GAAGCTGTGTATGGCTGTGCGAGTCTGTAATCACTGCATATTCG	1320					
Qy	1321	TGTGCTCAAGCGC	CACTCCCGTCTCGATAATGTTTTTGGCGCGACATCATATAACGGTT	1380					
Db	1321	TGTGCTCAAGCGC	CACTCCCGTCTCGATAATGTTTTTGGCGCGACATCATATAACGGTT	1380					
Qy	1381	CTGGCAAAATTTCT	GAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440					
Db	1381	CTGGCAAAATTTCT	GAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440					
Qy	1441	ATTGTGACCGATTA	CAATTTTCAACAGGAACACCGCAGTCGGTTTAGTGTTTTTCACGA	1500					
Db	1441	ATTGTGACCGATTA	CAATTTTCAACAGGAACACCGCAGTCGGTTTAGTGTTTTTCACGA	1500					
Qy	1501	GCACCTTCAACCA	AGGACCATAGATTATGAAACTGGAAGAGGTAACCTGTAATCTGG	1560					
Db	1501	GCACCTTCAACCA	AGGACCATAGATTATGAAACTGGAAGAGGTAACCTGTAATCTGG	1560					
Qy	1561	ATTAACCGCGATA	AAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620					
Db	1561	ATTAACCGCGATA	AAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620					
Qy	1621	ACCGGAATTAAGT	CACCGTTGAGCATCCGATAAATCTGGAAGAGAAATTCGCCAGGTT	1680					
Db	1621	ACCGGAATTAAGT	CACCGTTGAGCATCCGATAAATCTGGAAGAGAAATTCGCCAGGTT	1680					
Qy	1681	GGGCAACTGCGCA	TGGCCCTGACATTAATCTTCTGGGCACACCGCTTCTGCTGGCTAC	1740					
Db	1681	GGGCAACTGCGCA	TGGCCCTGACATTAATCTTCTGGGCACACCGCTTCTGCTGGCTAC	1740					
Qy	1741	GCTCAATCTGGC	CTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT	1800					
Db	1741	GCTCAATCTGGC	CTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT	1800					
Qy	1801	CGTTTTACTGGAT	GCGGTACGTTAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860					
Db	1801	CGTTTTACTGGAT	GCGGTACGTTAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860					
Qy	1861	GAAGCGTTATCG	CTGATTTATAAACAAGATCTGCTGCCGAAACCCCGCAAAACCTGGGAA	1920					
Db	1861	GAAGCGTTATCG	CTGATTTATAAACAAGATCTGCTGCCGAAACCCCGCAAAACCTGGGAA	1920					
Qy	1921	GAGATCCCGCG	CTGGATGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC	1980					
Db	1921	GAGATCCCGCG	CTGGATGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC	1980					
Qy	1981	CTGCAAGAACCG	TACTTCACTGGCCGCTGATTGCTGTGACCGGGGTTATCGGTTCAAG	2040					
Db	1981	CTGCAAGAACCG	TACTTCACTGGCCGCTGATTGCTGTGACCGGGGTTATCGGTTCAAG	2040					
Qy	2041	TATGAAAAACGG	CAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGCGAAAGCG	2100					
Db	2041	TATGAAAAACGG	CAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGCGAAAGCG	2100					

```
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAAAACAACACATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAAAACAACACATGAATGACACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGTTGG 2220
Db 2161 TCCATCGCAGAGCTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGTTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGFTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGFTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CGAACCAAGAGCTGGCAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CGAACCAAGAGCTGGCAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGTTATGCGCCACCATGGAAAAACGCCAGAAAGTGAAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATGCGCCACCATGGAAAAACGCCAGAAAGTGAAATCATG 2520
QY 2521 CGAACATCCCGCAGATCTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
Db 2521 CGAACATCCCGCAGATCTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGTCAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACACACAAATAACAAATAACAACTCGGGATCGAGGAAAGATTTCAGAAATTC 2700
Db 2641 AACACACACAAATAACAAATAACAACTCGGGATCGAGGAAAGATTTCAGAAATTC 2700
```

RESULT 14

```
US-10-263-153-20
; Sequence 20, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN.
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxo30del3C (52-300aa)
US-10-263-153-20
```

```
Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 CGGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CGGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCATTCAGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGCTCTTATACAGACCGTTTCCCGTGTGTGAACCAAGGCGAGCCAGCTTTCTCGGAAA 180
Db 121 GTGCTCTTATACAGACCGTTTCCCGTGTGTGAACCAAGGCGAGCCAGCTTTCTCGGAAA 180
QY 181 CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCACAAACCGCGTGGGAC 240
Db 181 CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCACAAACCGCGTGGGAC 240
QY 241 AACAACTGGGGGGCAACAGTCGTTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGGGGGCAACAGTCGTTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCCGTGCAGAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCCGTGCAGAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTAGAACGAAGCGCGTGAAGCCTGTAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTGTGATGTAGAACGAAGCGCGTGAAGCCTGTAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACCGTCAGTGGCTGATCAATTAATCTCGCTGGATGACACAGATGCCA 480
Db 421 TTCTCGCGCAACCGTCAGTGGCTGATCAATTAATCTCGCTGGATGACACAGATGCCA 480
QY 481 TTCTGTGGAGCTGCTGCACTAAATTTTCGCGGCTTAATTTCTTGATGTCTTCGACAGA 540
Db 481 TTCTGTGGAGCTGCTGCACTAAATTTTCGCGGCTTAATTTCTTGATGTCTTCGACAGA 540
QY 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGCTAGCGACTGGGGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGCTAGCGACTGGGGCTGGAGCATC 600
QY 601 TGGTGGATTTGGTCAACGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGGATTTGGTCAACGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGGCTCTGGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Db 661 CGGCTCTGGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
QY 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCCGGTTTTCACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCCGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTTCCACTGCGATGCTGGTTCACAAACGATCAGATGGCGCTGGGCGAA 840
Db 781 ATGAGGGCATGTTTCCACTGCGATGCTGGTTCACAAACGATCAGATGGCGCTGGGCGAA 840
QY 841 TGGCGCCATTACCGAGTCCGGGCTGGCGTGGTGGCGATATCTCGTAGTGGGATAGC 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTGGCGTGGTGGCGATATCTCGTAGTGGGATAGC 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAAGGATTTTC 960
QY 961 GCCTCTGGGCAAAACCGAGCTGGACCGTGTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGCAAAACCGAGCTGGACCGTGTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGTGTGGCCGCTCTCACTGGTGAAGAAAAAACCAACCTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGTGTGGCCGCTCTCACTGGTGAAGAAAAAACCAACCTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCCTCTCCCGCGGTTGGCGGATTATTAATGACGCTGGCAACGACAGTTT 1140
```

Db 1081 |||||CGAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCAAGTGGCAACAGAGTTT||| 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACGTGACGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACGTGACGGTGCAACCAATGCTTCTGGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATGCGCTGATATGTTGTTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATGCGCTGATATGTTGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTCACGA 1500
Qy 1501 GCACCTTCCACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCTGTAATCTGG 1560
Db 1501 GCACCTTCCACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCTGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAGTGGTGAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAGTGGTGAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATNAACCTGGAAGAGAAATTCACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATNAACCTGGAAGAGAAATTCACAGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCCGGAACAAAGCTTCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCCGGAACAAAGCTTCAGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATCCGCTAGCTTACACGCAAGCTGATGCTTACCGATCCTGTT 1860
Db 1801 CCGTTTACCTGGATCCGCTAGCTTACACGCAAGCTGATGCTTACCGATCCTGTT 1860
Qy 1861 GAAGGTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCTCGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTCGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGACGGGGTTATCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGACGGGGTTATCGTTCAAG 2040
Qy 2041 TATGAAACCGGCAAGTACGACATTAAGACGTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTAAGACGTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATGAATGACACACCGATTAC 2160
Qy 2161 TCCATCGGAAGCTGCTTAAATTAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220

Db 2161 TCCATCGGAAGCTGCTTAAATTAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACACAGCAAAAGTGAATTTATGGTTAAACGTTATCTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACACAGCAAAAGTGAATTTATGGTTAAACGTTATCTGCCGACCTTC 2280
Qy 2281 AAGGTCNAACCATCAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCNAACCATCAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCTG 2400
Qy 2401 GAAGCGGTTAATAAAGCAAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACAGAGAAAG 2460
Db 2401 GAAGCGGTTAATAAAGCAAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACAGAGAAAG 2460
Qy 2461 TTGGGGAAGATCCACGTTATTCGCGCCACCATGGAACGCCAGAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGGAAGATCCACGTTATTCGCGCCACCATGGAACGCCAGAAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGTCTG 2640
Qy 2641 AACCAACACAAATAACAATAACAACCAACCTCGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACCAACACAAATAACAATAACAACCAACCTCGGATCGAGGAGGATTCAGAAATTC 2700

RESULT 15
US-10-263-153-61
; Sequence 61, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxop30MIX1
US-10-263-153-61

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGCGATGATAGCCCGCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGCGATGATAGCCCGCGGAAGAGA 60
Qy 61 GTCAAATTCAGGTCGTGAATGTAACCAAGTAACTTATACGATGTCGACAGTATCGCG 120
Db 61 GTCAAATTCAGGTCGTGAATGTAACCAAGTAACTTATACGATGTCGACAGTATCGCG 120

QY	121	GTGTCCTCTTATCAGACGGTTTCCGCGTGGTGAACCAAGGCGACGAGCTTTCTGCGAAA	180
Db	121	GTGTCCTCTTATCAGACGGTTTCCGCGTGGTGAACCAAGGCGACGAGCTTTCTGCGAAA	180
QY	181	CGCGGAAAAGTGGAGCGCGGATGGCGAGCTGAATATCATTCCTCAAACCGCGTGGCAC	240
Db	181	CGCGGAAAAGTGGAGCGCGGATGGCGAGCTGAATATCATTCCTCAAACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCAACACAGTCGTGCTGATTGGCGTTGCGCACTCCAGCTTGGCCCTGC	300
Db	241	AACAACTGGCGGCAACACAGTCGTGCTGATTGGCGTTGCGCACTCCAGCTTGGCCCTGC	300
QY	301	ACGCGCGTGCMAATTTGTCGGCGGCAATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCGTGCMAATTTGTCGGCGGCAATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
QY	361	TGGTGGTGTGATGGTAGAACGAGCGGCGTGAAGCCGTGAAGCGGCGTGCACAATC	420
Db	361	TGGTGGTGTGATGGTAGAACGAGCGGCGTGAAGCCGTGAAGCGGCGTGCACAATC	420
QY	421	TTCTCGGCAACGCGTCAGTGGGCTGATCATTTAACTCTCGCGCGATCAACTGGGTGCCA	480
Db	421	TTCTCGGCAACGCGTCAGTGGGCTGATCATTTAACTCTCGCGCGATCAACTGGGTGCCA	480
QY	481	TTGCTGTGAAGCTGCCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540
Db	481	TTGCTGTGAAGCTGCCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTAGCGGCTGAGCGGCGTGGAGATC	600
Db	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTAGCGGCTGAGCGGCGTGGAGATC	600
QY	601	TGGTGCATTTGGGTCAACGACMAATTCGGCTGTTAGCGGCGCCATTAAGTTCTGTCTCGG	660
Db	601	TGGTGCATTTGGGTCAACGACMAATTCGGCTGTTAGCGGCGCCATTAAGTTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCGCGATG	720
Db	661	CGCGTCTGCGTCTGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCGCGATG	720
QY	721	CGGAAACGGGAAGGACATGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAAACGGGAAGGACATGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGGCATGTTTCCCACTCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGCAA	840
Db	781	ATGAGGGCATGTTTCCCACTCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGCGCATTTACCGAGTTCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACCGAGTTCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAGACAGCTCATGTTATATCCGCGGTTAACCAATCAAAACAGGATTTTC	960
Db	901	ACGATACCGAGACAGCTCATGTTATATCCGCGGTTAACCAATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGGCAATAGCTGTTGCCGCTCTACTGTGTAAGAAAACCAACCTCGCGCCCAATA	1080
Db	1021	AGGGCAATAGCTGTTGCCGCTCTACTGTGTAAGAAAACCAACCTCGCGCCCAATA	1080
QY	1081	CGAAAACCGGCTCTCCCGCGGTTGGCCGATTCATTATGACGCTGGCACGACGAGTTT	1140
Db	1081	CGAAAACCGGCTCTCCCGCGGTTGGCCGATTCATTATGACGCTGGCACGACGAGTTT	1140
QY	1141	CCGACCTGGAAGCGGCGAGTGACGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGACCTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200

QY	1201	GCACAAATTTCTCATGTTTGCACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATTTCTCATGTTTGCACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCCATCCGAAAGCTGTGTATGGCTGTGCAGGTCGTAATCATCTGCATAAATTCG	1320
Db	1261	TCAGGACGCCATCCGAAAGCTGTGTATGGCTGTGCAGGTCGTAATCATCTGCATAAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCATACGGTT	1380
QY	1381	CTGCGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCTGGCTCGTAATGTGTGGA	1440
Db	1381	CTGCGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCTGGCTCGTAATGTGTGGA	1440
QY	1441	ATTGTGAGCGGATAAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAAATCATCTGGCT	1500
Db	1441	ATTGTGAGCGGATAAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAAATCATCTGGCT	1500
QY	1501	GCATTTACCAACAAAGGACCATAGATTTATGAAAACCTGAAGAGGTAAACTGGTAATCTGG	1560
Db	1501	GCATTTACCAACAAAGGACCATAGATTTATGAAAACCTGAAGAGGTAAACTGGTAATCTGG	1560
QY	1561	ATTAAACGGCGATAAAGGCTATTAACGGTCTGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATTAACGGTCTGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGATCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
Db	1621	ACCGGAATTAAGATCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
QY	1681	GGCGCAACTTGGCGATCGCCCTGACATTTCTCTGGGCAACAGCCGCTTGGTGCTTAC	1740
Db	1681	GGCGCAACTTGGCGATCGCCCTGACATTTCTCTGGGCAACAGCCGCTTGGTGCTTAC	1740
QY	1741	GCTCAATCTGGCCTGTGCTGAAATCACCCCGGACAAAGCGTTCCAGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCTGTGCTGAAATCACCCCGGACAAAGCGTTCCAGACAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGCCGTAAGTTTAAACGGCAAGCTGATTTGCTTACCCTGCTGTT	1860
Db	1801	CGGTTTACCTGGGATGCCGTAAGTTTAAACGGCAAGCTGATTTGCTTACCCTGCTGTT	1860
QY	1861	GAAGCGTATTCGCTGATTTTATAACAAAGATCTGCTGCGCAACCCGCCAAAACCTGGGA	1920
Db	1861	GAAGCGTATTCGCTGATTTTATAACAAAGATCTGCTGCGCAACCCGCCAAAACCTGGGA	1920
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCCCTGGCGGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCCCTGGCGGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAACGGCAAGTACGACATTAAGACCGTGGCGGTGATTAACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAAACGGCAAGTACGACATTAAGACCGTGGCGGTGATTAACGCTGGCGCGAAAGCG	2100
QY	2101	GCTCTGACCTTCTGGTTGACCTGATTAATAACAAACATGAATGACAGACCCGATTTAC	2160
Db	2101	GCTCTGACCTTCTGGTTGACCTGATTAATAACAAACATGAATGACAGACCCGATTTAC	2160
QY	2161	TCCATCGCAGAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAGAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
QY	2221	GCATGGTCCACATCGACACCGCAAGTGAATTTATGGTGAACGGTACTGCCACCTTC	2280
Db	2221	GCATGGTCCACATCGACACCGCAAGTGAATTTATGGTGAACGGTACTGCCACCTTC	2280
QY	2281	AAGGCTCAACCATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAACCGCGCCAGT	2340

Db	2281	AAGGGTCAACCATCAAAACCGTTCTGTTGGGTGCTGAGCGCAGGTATTAACGCGCCAGT	2340
Qy	2341	CCGAAACAAAGAGCTGGCAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCG	2400
Qy	2401	GAAGCGGTTAATAAAGACAAACCCCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCCCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGCGGAAAGATCCAACGTATTGCGCGCACCATGTGAAAAACGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGCGGAAAGATCCAACGTATTGCGCGCACCATGTGAAAAACGCCAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCGCGAGATGTCGCTTTTCTGGTATGCCGTGCGTACTCGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCGCGAGATGTCGCTTTTCTGGTATGCCGTGCGTACTCGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGCTGAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGGTCG	2640
Db	2581	GCCAGCGGTGCTGAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGGTCG	2640
Qy	2641	AACAACAAACAATAACAATAACAACACTCCGGATTCGAGGGAAGGATTTCAGAAATTC	2700
Db	2641	AACAACAAACAATAACAATAACAACACTCCGGATTCGAGGGAAGGATTTCAGAAATTC	2700

Search completed: August 3, 2005, 04:22:07
Job time : 1316.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-17
Perfect score: 3300
Sequence: 1 ccgacacatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	23.6	794	5	BQ751936
2	770.6	23.4	779	7	CK394523
3	768	23.3	769	7	CK118014
4	762	23.1	793	5	BQ751655
5	755.4	22.9	806	5	BQ751220
6	684.8	20.8	689	6	CB863541
7	676.4	20.5	712	1	AL038548
8	676	20.6	710	8	AY080106
9	647.4	19.6	660	1	AL044483
10	641	19.4	641	4	BJ063872
11	615.8	18.7	624	6	CB863814
12	612	18.5	794	1	AL045353
13	603	18.3	752	9	CL655575
14	589.8	17.9	608	8	AY080094
15	586.4	17.8	589	1	AL037742
16	581.2	17.6	601	1	AV594538
17	579	17.5	579	1	AL043868
18	570	17.3	782	7	CK394246
19	569	17.2	774	1	AL045337
20	563.2	17.1	585	9	CL658017
21	562.6	17.0	723	1	AL039416
22	549.4	16.6	753	1	AL045341
23	546.8	16.6	782	1	AL037051
24	545.4	16.5	547	6	CA895956

C 25	545	16.5	648	7	CK394349
C 26	540.2	16.4	569	1	AL039077
C 27	534.4	16.2	644	6	CB862524
C 28	534	16.2	734	1	AL048427
C 29	529	16.0	740	1	AL042909
C 30	528	16.0	720	6	CA881974
C 31	523.2	15.9	528	1	AL038811
C 32	519.4	15.7	521	6	CA890032
C 33	519	15.7	607	6	CB862099
C 34	516.4	15.6	518	6	CA886082
C 35	510.2	15.5	686	1	AL044407
C 36	509.4	15.4	511	6	CA895436
C 37	498	15.1	498	1	AL039076
C 38	494.8	15.0	1025	1	AL038025
C 39	487.6	14.8	498	7	CK394397
C 40	477.6	14.5	491	6	CB862715
C 41	465.8	14.1	679	1	AL039128
C 42	462.8	14.0	490	1	AL039589
C 43	458	13.9	458	6	CA887583
C 44	456.4	13.8	471	1	AL039649
C 45	449.2	13.6	612	6	CB863409

ALIGNMENTS

RESULT 1
LOCUS BQ751936 794 bp mRNA linear EST 18-JUL-2002
DEFINITION EST632499 DSCT Colletotrichum trifolii cDNA clone pDSCT8-67, mRNA
sequence.
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 794)
AUTHORS Samad D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
Cheung, F. and Fraser, C.M.
ESTS from mycelia of Colletotrichum trifolii race 1
Other ESTs: EST632498
Unpublished (2002)
CONTACT: Deborah A. Samad
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debby@puccini.crl.umn.edu
TIGR sequence name: MTSAH67TV More information is available at:
www.medicago.org
Seq primer: (gtA AtA CgA Ctc Act AtA ggg C).
Location/Qualifiers
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSCT8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"

FEATURES
source

1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSCT8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"

/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl1 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce.

E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

```
Query Match      23.6%; Score 778; DB 5; Length 794;
Best Local Similarity 98.7%; Pred. No. 6e-218;
Matches 784; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 GACACCATCGAATGTCGCAAAACCTTTCGGGTATGCGATGATAGCGCCGGAAGAGAGT 62
DB 1 GACACCTTTTGAATGCGCAAAACCTTTCGGGTATGCGATGATAGCGCCGGAAGAGAGT 60

QY 63 CAATTTCAGGTCGTCGTAATGTAACACAGTAACTTATACGATGTCGACAGATGTCGGT 122
DB 61 CAATTTCAGGTCGTCGTAATGTAACACAGTAACTTATACGATGTCGACAGATGTCGGT 120

QY 123 GTCTCTTATCAGACCGTTTCCCGCTGCTGTAACACAGGCGCCAGCTTTCTGCGAAAACG 182
DB 121 GTCTCTTATCAGACCGTTTCCCGCTGCTGTAACACAGGCGCCAGCTTTCTGCGAAAACG 180

QY 183 CGGGAAGAGTGGAGCGCGATGCGGAGCTGAATTAACATTCACACCGCGTGCGACAA 242
DB 181 CGGGAAGAGTGGAGCGCGATGCGGAGCTGAATTAACATTCACACCGCGTGCGACAA 240

QY 243 CAACCTGGCGGCAACACAGTCGTGCTGATGCGGCTTGCCACTCCAGTCTGCGCCCTGCAC 302
DB 241 CAACCTGGCGGCAACACAGTCGTGCTGATGCGGCTTGCCACTCCAGTCTGCGCCCTGCAC 300

QY 303 GCGCGTCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCAATCGGTCGCCAGCGTG 362
DB 301 GCGCGTCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCAATCGGTCGCCAGCGTG 360

QY 363 GTGGTGTGATGTAAGACGAGCGGCTGGAAGCCTGTAAGCGGCGGTCGCAATCTT 422
DB 361 GTGGTGTGATGTAAGACGAGCGGCTGGAAGCCTGTAAGCGGCGGTCGCAATCTT 420

QY 423 CTCGCGCAACCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCGAGATGCCATT 482
DB 421 CTCGCGCAACCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCGAGATGCCATT 480

QY 483 GCTGTGAGCTGCTGACATAATGTTCCGCGGTTATTTCTTGATGTCCTGACAGACA 542
DB 481 GCTGTGAGCTGCTGACATAATGTTCCGCGGTTATTTCTTGATGTCCTGACAGACA 540

QY 543 CCCATCAACAGTATTATTTCTCCATGAAGACGCTAGCGACTGGGCGTGGAGCATCTG 602
DB 541 CCCATCAACAGTATTATTTCTCCATGAAGACGCTAGCGACTGGGCGTGGAGCATCTG 600

QY 603 GTCGATTTGGTTCACAGCAAAATCGCGTGTGTAGCGGCGCCATTAAAGTTCTGTCGCGG 662
DB 601 GTCGATTTGGTTCACAGCAAAATCGCGTGTGTAGCGGCGCCATTAAAGTTCTGTCGCGG 660

QY 663 GCTGTGGTCTGGTGGTCGATTAATATCTCTGATCAATCAAAATTCAGCCGATAGCG 722
DB 661 GCTGTGGTCTGGTGGTCGATTAATATCTCTGATCAATCAAAATTCAGCCGATAGCG 720

QY 723 GAACGGGAAGCGACTGAGTGGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGAAT 782
DB 721 GAACGGGAAGCGACTGAGTGGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGAAT 780

QY 783 GAGGGCATCGTTTC 796
DB 781 GAGGGCATCGTTTC 794
```

RESULT 2
CK394523/c
LOCUS hggada408 Gland Cell Amplified cDNA Library Heterodera glycines
DEFINITION hggada408 Gland Cell Amplified cDNA Library Heterodera glycines

cDNA, mRNA sequence.
CK394523
VERSION CK394523.1 GI:40389794
KEYWORDS EST.
SOURCE Heterodera glycines
ORGANISM Heterodera glycines
REFERENCE Heterodera glycines
AUTHORS Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
1 (bases 1 to 779)
Wang, X., Allen, R., Gao, B., Goellner, M., Maier, T., Baum, T.,
Hussey, R. and Davis, E.
Submission: Wang, Allen, Gao, Goellner, Maier, Baum, Hussey, Davis
Unpublished (2003)
Contact: Tom Maier
Department of Plant Pathology, Baum Lab
Iowa State University
351 Bessey Hall, Ames, IA 50011, USA
Tel: 515-294-8854
Fax: 515-294-9420
Email: tmaier@iastate.edu
Heterodera glycines Gland Cell Amplified cDNA Library, single pass
sequence.

FEATURES

source

1..779
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell Amplified cDNA Library"
/note="Organ: gland cell; Vector: pSportII"

ORIGIN

```
Query Match      23.4%; Score 770.6; DB 7; Length 779;
Best Local Similarity 99.4%; Pred. No. 9.3e-216;
Matches 773; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 87 CCAGTAACGTTATACGATGTCGAGAGATGCGCGGTGTCCTTTATCAGACCGTTTCCCGC 146
DB 779 CCAGTAACGTTATACGATGTCGAGAGATGCGCGGTGTCCTTTATCAGACCGTTTCCCGC 720

QY 147 GTGGTGAACAGCGCCAGCCAGCTTTCTCGGAAAACCGGGAAAAAGTGAAGCGCGGATG 206
DB 719 GTGGTGAACAGCGCCAGCCAGCTTTCTCGGAAAACCGGGAAAAAGTGAAGCGCGGATG 660

QY 207 GCGGAGCTGAATTAACATTCACACCGCGTGGCAACAACCTGCGGGCAACAGTCTGTTG 266
DB 659 GCGGAGCTGAATTAACATTCACACCGCGTGGCAACAACCTGCGGGCAACAGTCTGTTG 600

QY 267 CTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGTCACGCGCGCTCGCAAAATTTGCGGGCG 326
DB 599 CTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGTCACGCGCGCTCGCAAAATTTGCGGGCG 540

QY 327 ATTAATCTCGGCCCATCAACTGGGTGCGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 386
DB 539 ATTAATCTCGGCCCATCAACTGGGTGCGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 480

QY 387 GGCGTCCAAAGCCTGTAAGCGCGGTGCAATCTTCTCGCGCAACCGCTGAGTGGGCTG 446
DB 479 GGCGTCCAAAGCCTGTAAGCGCGGTGCAATCTTCTCGCGCAACCGCTGAGTGGGCTG 420

QY 447 ATCAATTAATCTCGGTGGATGACAGGATGCCATTTGCTGTGGAACTGCTGCTCACTAAT 506
DB 419 ATCAATTAATCTCGGTGGATGACAGGATGCCATTTGCTGTGGAACTGCTGCTCACTAAT 360

QY 507 GTTCCGGCGTTATTTCTTGTATGTCCTGACCGACACCCATCAACAGTATTATTTCTCC 566
DB 359 GTTCCGGCGTTATTTCTTGTATGTCCTGACCGACACCCATCAACAGTATTATTTCTCC 300

QY 567 CATGAAGACGGTACGCGACTGGCGGTGGAGCATCTGGTGGCATTTGGGTCAACAGCAATC 626
DB 299 CATGAAGACGGTACGCGACTGGCGGTGGAGCATCTGGTGGCATTTGGGTCAACAGCAATC 240
```


COMMENT

Other_ESTs: EST632217
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org

FEATURES

Location/Qualifiers
 1..793
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST6-59"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSTCT"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 23.1%; Score 762; DB 5; Length 793;
 Best Local Similarity 99.1%; Pred. No. 3.3e-213;
 Matches 777; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

3 GACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATATAGCGCCGGAAGAGAGT 62
 Db 1 GACATCTTCGAATGGCGCAAAACCTTTCGCGGTATGGCATATAGCGCCGGAAGAGAGT 60

63 CAATTCAGGGTGGTGAATGTGAACACCAAGTATACGATGTCGAGAGTATGCCGGT 122
 Db 61 CAATTCAGGGTGGTGAATGTGAACACCAAGTATACGATGTCGAGAGTATGCCGGT 120

123 GTCTCTTATCAGACCGTTTCCGCGGTGGTGAACAGGCGCAGCCACGTTTCTGCGAAAACG 182
 Db 121 GTCTCTTATCAGACCGTTTCCGCGGTGGTGAACAGGCGCAGCCACGTTTCTGCGAAAACG 180

183 CGGGAAAAAGTGAAGCGGCATCGCGAGCTGAATTCACATTCCTCCACCGGTGGCAAA 242
 Db 181 CGGGAAAAAGTGAAGCGGCATCGCGAGCTGAATTCACATTCCTCCACCGGTGGCAAA 240

243 CAATCGCGGCGCAACAGTGTCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 302
 Db 241 CAATCGCGGCGCAACAGTGTCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 300

303 CGCGCGTGCAGAAATGTTCGCGGCGATTAATCTCGCCCGCATCACTGGGTGCCAGGTTG 362
 Db 301 CGCGCGTGCAGAAATGTTCGCGGCGATTAATCTCGCCCGCATCACTGGGTGCCAGGTTG 360

363 GTGGTGTGATGGTAGAACAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACATCTT 422
 Db 361 GTGGTGTGATGGTAGAACAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACATCTT 420

423 CTGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCCGTGGATGACCAAGATGCCATT 482
 Db 421 CTGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCCGTGGATGACCAAGATGCCATT 480

QY 483 GCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTTGACAGACA 542
 Db |||||
 QY 481 GCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTTGACAGACA 540
 Db |||||
 QY 543 CCCATCAACAGTATTATTTCTCCATGAAGACGGTACGGACTGGGGTGGAGCATCTG 602
 Db |||||
 QY 541 CCCATCAACAGTATTATTTCTCCATGAAGACGGTACGGACTGGGGTGGAGCATCTG 600
 Db |||||
 QY 603 GTCGCATTGGTTCACCAAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTCTCGCGG 662
 Db |||||
 QY 601 GTCGCATTGGTTCACCAAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTCTCGCGG 660
 Db |||||
 QY 663 CGTCTGCGTCTGCGTGGCTGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
 Db |||||
 QY 661 CGTCTGCGTCTGCGTGGCTGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720
 Db |||||
 QY 723 GAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATCTGTAAT 782
 Db |||||
 QY 721 GAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATCTGTAAT 778
 Db |||||
 QY 783 GAGG 786
 Db |||||
 QY 779 GAGG 782

RESULT 5
 BQ751220 806 bp mRNA linear EST 18-JUL-2002
 LOCUS EST631783 DSCT Colletotrichum trifolii cDNA clone pDST3-64, mRNA
 DEFINITION sequence.
 ACCESSION BQ751220
 VERSION BQ751220.1 GI:21906625
 KEYWORDS EST
 SOURCE Colletotrichum trifolii
 ORGANISM Colletotrichum trifolii

REFERENCE
 AUTHORS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
 mitosporic Phyllachoraceae; Colletotrichum.
 TITLE 1 (bases 1 to 806)
 JOURNAL Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
 COMMENT ESTs from mycelia of Colletotrichum trifolii race 1
 Unpublished (2002)
 Other ESTs: EST631782

Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org

FEATURES

Location/Qualifiers
 1..806
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST3-64"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSTCT"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid

lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK⁺. Aliquots of the ligation were used to transform *E. coli* Disalpa which were plated onto medium with X-gal for selection of recombinants."

ORIGIN	Query Match Best Local Similarity Matches	22.9%; 97.4%; 790;	Score 755.4; Pred. No. 2.9e-211; 0;	DB 5; Length 806; Mismatches 16;	Indels 5;	Gaps 2
Qy	3	GACACCATCGAATCGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCGCGAAGAGAGT	62			
Db	1	GACACTTTGCAATGGCGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCGCGAAGAGAGT	60			
Qy	63	CAATTTCAGGGTGGTGAATGTGCAAAACCAAGTAACTGTTATACGATGTCGCAGAGTATGCCCGT	122			
Db	61	CAATTTCAGGGTGGTGAATGTGAAACCAAGTAACTGTTATACGATGTCGCAGAGTATGCCCGT	120			
Qy	123	GTCTCTTATCAGACCGTTTTCGCGGTGTTGAACAGGCGCAGCCAGCTTTTCTGCGAAAACG	182			
Db	121	GTCTCTTATCAGACCGTTTTCGCGGTGTTGAACAGGCGCAGCCAGCTTTTCTGCGAAAACG	180			
Qy	183	CGGGAAAAAGTGGAAAGCGGCGGATGGCGAGCTGAATTAATTCATTCCTCAAACGGCGTGCGACAA	242			
Db	181	CGGGAAAAAGTGGAAAGCGGCGGATGGCGAGCTGAATTAATTCATTCCTCAAACGGCGTGCGACAA	240			
Qy	243	CNACTGGCGGCAACACAGTCTGCTGATTCGGCTGGCGATGGCCACTCCAGTCTGGCCCTGCAC	302			
Db	241	CNACTGGCGGCAACACAGTCTGCTGATTCGGCTGGCGATGGCCACTCCAGTCTGGCCCTGCAC	300			
Qy	303	GGCGCGTGCATAATTTGTCGGCGGCAATTAATCTTCGCGCGGATCAACTGGGTGCCAGCGTG	362			
Db	301	GGCGCGTGCATAATTTGTCGGCGGCAATTAATCTTCGCGCGGATCAACTGGGTGCCAGCGTG	360			
Qy	363	GTGGTGTGATGTTAGAACGAAGCGGCGTTCGAAGCCTCTGAAGCGGCGGTGCACAACTCTT	422			
Db	361	GTGGTGTGATGTTAGAACGAAGCGGCGTTCGAAGCCTCTGAAGCGGCGGTGCACAACTCTT	420			
Qy	423	CTCGCGCAACCGGTCAGTGGGCTGATCAATTAACATATCCGCTGGATGACAGGATGCCATT	482			
Db	421	CTCGCGCAACCGGTCAGTGGGCTGATCAATTAACATATCCGCTGGATGACAGGATGCCATT	480			
Qy	483	GCTGTGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACA	542			
Db	481	GCTGTGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACA	540			
Qy	543	CCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCACTGGGCGTGAGGATCTG	602			
Db	541	CCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCACTGGGCGTGAGGATCTG	600			
Qy	603	GTCCGATTTGGTTCACCAAGCAATTCGCGCTGTGTAGCGGGCCCATTAAGTCTGTCTCGGG	662			
Db	601	GTCCGATTTGGTTCACCAAGCAATTCGCGCTGTGTAGCGGGCCCATTAAGTCTGTCTCGGG	660			
Qy	663	CGTCTGGTCTGGCTGGCTGCAATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGG	722			
Db	661	CGTCTGGTCTGGCTGGCTGCAATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGG	720			
Qy	723	GAAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTTCAACCAACCATGCGAAATGCTGAAT	782			
Db	721	GAAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTTCAACCAACCATGCGAAATGCTGAAT	777			
Qy	783	GAGGCGATCGTTCCCACTGCGAATGCTGGTTG	813			
Db	778	GAGG--CATCGTTCCACTGCGAATGCTGGTTG	806			

RESULT 6					
CB863541/c					
LOCUS	CB863541	689 bp	mRNA	linear	EST 22-APR-2003
DEFINITION	HH04A08y HH Hordeum vulgare cDNA clone HH04A08 3-PRIME, mRNA sequence.				


```

QY 441 GGGCTGATCATTAATATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCCTGC 500
Db |||||
QY 269 GGGCTGATCATTAATATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCCTGC 210
Db |||||
QY 501 ACTAATGTTCCGGCGTATTCTTCTGATGCTCTGACACAGACACCCATCAACAGTATTATT 560
Db |||||
QY 209 ACTAATGTTCCGGCGTATTCTTCTGATGCTCTGACACAGACACCCATCAACAGTATTATT 150
Db |||||
QY 561 TTCTCCCATGAAGACGGTACCGCATCGGCGTGGAGCATCTGGTCGCATTGGGTCAACAG 620
Db |||||
QY 149 TTCTCCCATGAAGACGGTACCGCATCGGCGTGGAGCATCTGGTCGCATTGGGTCAACAG 90
Db |||||
QY 621 CAAATCGCGCTGTTAGCGGCGCCATTAAGTCTGCTCGGCGCGTCTCGTCTGCGCTGCG 680
Db |||||
QY 89 CAAATCGCGCTGTTAGCGGCGCCATTAAGTCTGCTCGGCGCGTCTCGTCTGCGCTGCG 30
Db |||||
QY 681 TGGCATAAATATCTCACATCGCAATCAAT 709
Db |||||
QY 29 TGGCATAAATATCTCACATCGCAATCAAT 1

RESULT 7
AL038548 712 bp mRNA linear EST 04-SEP-2003
LOCUS DKZP566E0346_r1 566 (synonym: hfkd2) Homo sapiens cDNA clone
DEFINITION DKZP566E0346 5', mRNA sequence.
ACCESSION AL038548
VERSION AL038548.1 GI:5407738
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 712)
AUTHORS Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MedGenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZp566E0346) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: Clone@rzpd.de.
FEATURES
source
1..712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp566E0346"
/tissue_type="kidney"
/dev_stage="fetal"
/lab_host="X1-2blue"
/clone_lib="566 (synonym: hfkd2)"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match 20.5%; Score 676.4; DB 1; Length 712;
Best Local Similarity 99.6%; Pred. No. 6.4e-188;
Matches 699; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 33 GGTATGGCATGATAGCCCGGAGAGAGTCAATTACAGGTGGTGTGATGTGAACAGTA 92
Db 712 GGTATGGCATGATAGCCCGGAGAGAGTCAATTACAGGTGGTGTGATGTGAACAGTA 653
QY 93 ACGTTATACAGTGTGCGAGATGTCGGGTGTCTCTTATCAGACCGGTTCCCGCGTGGTG 152
, |||||

```

```

Db 652 ACGTTATACAGTGTGCGAGATGTCGGGTGTCTCTTATCAGACCGGTTTCCCGCGTGGTG 593
QY 153 AACAGCGCAGCCACGTTTCTCGAAACCGCGG - AAAAAAGTGAACCGCGATGGCGGA 211
Db |||||
QY 592 AACAGCGCAGCCACGTTTCTCGAAACCGCGGAAAAAGTGAACCGCGATGGCGGA 533
Db |||||
QY 212 GCTCAATTTACATTCCTCCAAACCGCGTGGCACAACAACTCGCGGGCAACAGTCTGCTGAT 271
Db |||||
QY 532 GCTGAATTAATTCCTCCAAACCGCGTGGCACAACAACTCGCGGGCAACAGTCTGCTGAT 473
Db |||||
QY 272 TGGCGTTGCCACCTCCAGTCTGGCCCTGCGAC - GCGCCGTGCGCAAAATTTGTGCGCGCGATTA 330
Db |||||
QY 472 TGGCGTTGCCACCTCCAGTCTGGCCCTGCGACGCGCCGTGCGCAAAATTTGTGCGCGCGATTA 413
QY 331 AATCTCGCGCGCATCAACTGGGTGCCAGCGTGGTGTGTCGATGTTAGAACGAAACGCGCG 390
Db |||||
QY 412 AATCTCGCGCGCATCAACTGGGTGCCAGCGTGGTGTGTCGATGTTAGAACGAAACGCGCG 353
Db |||||
QY 391 TCGAAGGCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACCGGTCAAGTGGGCTGATCA 450
Db |||||
QY 352 TCGAAGGCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACCGGTCAAGTGGGCTGATCA 293
QY 451 TTAATACTACCGCTGGATGACACAGATGCCATTGCTGTGGAAGCTGCGCTGCACATAATGTTTC 510
Db |||||
QY 292 TTAATACTACCGCTGGATGACACAGATGCCATTGCTGTGGAAGCTGCGCTGCACATAATGTTTC 233
QY 511 CGGCGTTATTCTTGTATGTCCTGACGACACACCCATCAACAGTATTATTTTCTCCCATG 570
Db |||||
QY 232 CGGCGTTATTCTTGTATGTCCTGACGACACACCCATCAACAGTATTATTTTCTCCCATG 173
QY 571 AAGACGCTACGCGATGGCGGTGGAGCATCTGCTCGCATTTGGGTCAACCAAAATCGCGC 630
Db |||||
QY 172 AAGACGCTACGCGATGGCGGTGGAGCATCTGCTCGCATTTGGGTCAACCAAAATCGCGC 113
QY 631 TGTAGCGGCGCCATTAAAGTTCTGCTCGCGCGTCTGGGTCTGGGTGGTGGCTGCATAAAT 690
Db |||||
QY 112 TGTAGCGGCGCCATTAAAGTTCTGCTCGCGCGTCTGGGTCTGGGTGGTGGCTGCATAAAT 53
QY 691 ATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGNAAG 732
Db |||||
QY 52 ATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGNAAG 11

RESULT 8
AY080106 710 bp DNA linear GSS 06-NOV-2002
LOCUS AY080106
DEFINITION AY080106 Scripps Pier (La Jolla, CA) uncultured virus community
uncultured marine virus genomic clone SI05lp3G6L, genomic survey
sequence.
ACCESSION AY080106
VERSION AY080106
KEYWORDS GSS.
SOURCE uncultured marine virus
ORGANISM uncultured marine virus
REFERENCE 1 (bases 1 to 710)
AUTHORS Viruses; environmental samples.
Mead,D., Azam,P. and Rohwer,F.
Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
Genomic analysis of uncultured marine viral communities
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
TITLE JOURNAL
MEDLINE 22294988
PUBMED 12384570
COMMENT Contact: Rohwer F
Biocgy Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@unstroke.sdsu.edu
Class: shotgun.
FEATURES
Location/Qualifiers
1..710
/organism="uncultured marine virus"
source

```

```
/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone="SI0513P36L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
community"
/notes="Marine viruses were isolated from 200 liters of
surface seawater using a combination of differential
filtration and density-dependent gradient centrifugation.
Linker-amplified shotgun libraries were constructed by
randomly shearing the total marine viral community DNA,
end-repairing, ligating dsDNA linkers to the ends, and
amplifying the fragments using Vent DNA polymerase. The
resulting fragments were ligated into the pSMART vector
and electroporated into MC12 cells (Lucigen; Middleton,
WI)"

ORIGIN
Query Match      20.5%; Score 676; DB 8; Length 710;
Best Local Similarity 98.9%; Pred. No. 8.4e-188;
Matches 590; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 400 GTAAAGCGCGGTGCACAATCTTCTCGCGCAACCGCTCAGTGGGTGATCAATTAACATATC 459
Db      |||
Qy 460 CGCTGGATGACCGATGCCATTGCTGTGGAAGCTGCTGCACATATGTTCCGGCGTTAT 519
Db      |||
Qy 68  CGCTGGATGACCGATGCCATTGCTGTGGAAGCTGCTGCACATATGTTCCGGCGTTAT 127
Db      |||
Qy 520 TTCTTGATGTCTTGACGACACCCATCAACAGTATTATTTCTCCCATGAAGACGTA 579
Db      |||
Qy 128 TTCTTGATGTCTTGACGACACCCATCAACAGTATTATTTCTCCCATGAAGACGTA 187
Db      |||
Qy 580 CGCGACTGGCGTGAGCATCTGTCGATTTGGGTCAACAGAAATCGCGTGTAGCGG 639
Db      |||
Qy 188 CGCGACTGGCGTGAGCATCTGTCGATTTGGGTCAACAGAAATCGCGTGTAGCGG 247
Db      |||
Qy 640 GCCCATTAAGTCTGTCTCGGGCGTCTGCTGGCTGGCTGGCATAAATATCTCACTC 699
Db      |||
Qy 248 GCCCATTAAGTCTGTCTCGGGCGTCTGCTGGCTGGCTGGCATAAATATCTCACTC 307
Db      |||
Qy 700 GCAATCAAAATTCAGCGGATAGCGGAACGGAAGGAGTGGAGTGCCTATGTCGGGTTTC 759
Db      |||
Qy 308 GCAATCAAAATTCAGCGGATAGCGGAACGGAAGGAGTGGAGTGCCTATGTCGGGTTTC 367
Db      |||
Qy 760 AACAAACCATGCAATGCTGAATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACG 819
Db      |||
Qy 368 AACAAACCATGCAATGCTGAATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACG 427
Db      |||
Qy 820 ATCAGATGGCGTGGCGCAATTCGCGGCATTTACCGAGTCCGGGTGCGGTGTCGGG 879
Db      |||
Qy 428 ATCAGATGGCGTGGCGCAATTCGCGGCATTTACCGAGTCCGGGTGCGGTGTCGGG 487
Db      |||
Qy 880 ATATCTCGGTAGTGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAA 939
Db      |||
Qy 488 ATATCTCGGTAGTGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAA 547
Db      |||
Qy 940 CCACCATCAACAGGATTTTCGCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAAC 999
Db      |||
Qy 548 CCACCATCAACAGGATTTTCGCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAAC 607
Db      |||
Qy 1000 TCTCTCAGGGCGAGCGGTGAAGGCAATTCAGTCTTTCGCGTCTCACTGGTGAAGAA 1059
Db      |||
Qy 608 TCTCTCAGGGCGAGCGGTGAAGGCAATTCAGTCTTTCGCGTCTCACTGGTGAAGAA-AA 666
Db      |||
Qy 1060 AAACCCACCTGGCGCCCAATAGCAAAACCGCTCTCC 1097
Db      |||
Qy 667 AAACCCACCTGGCGCCCAATAGCAAAACCGNCTTTCCC 704
Db      |||
```

RESULT 9
AL044483/c
LOCUS AL044483 660 bp mRNA linear EST 04-SEP-2003

```
DEFINITION DKF2p4341102.s1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKF2p4341102.3, mRNA sequence.
VERSION AL044483
KEYWORDS EST.
SOURCE AL044483.1 GI:5432701
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Ansorge, W., Benes, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and
Wismann, S.
TITLE EST (Ansorge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wismann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wismann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKF2p4341102) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..660
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p4341102"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match      19.6%; Score 647.4; DB 1; Length 660;
Best Local Similarity 99.7%; Pred. No. 2.4e-179;
Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 85 AACAGTAACGTTATACGATTCGACAGTATGCGGTGTTCTTTATCAGACCGTTCCC 144
Db      |||
Qy 660 AACAGTAACGTTATACGATTCGACAGTATGCGGTGTTCTTTATCAGACCGTTCCC 601
Db      |||
Qy 145 CGGTGGTAACAGCCAGCCACGTTTCTGCGAAACGCGGAAAGTGGAGCGCGGA 204
Db      |||
Qy 600 CGGTGGTGAACAGCCAGCCAGCCACGTTTCTGCGAAACGCGGAAAGTGGAGCGCGGA 541
Db      |||
Qy 205 TGGCGGAGCTCAATTACATTCCTCCCAACCGCGTGGCAACAACCTGGCGGCAACAGTCGT 264
Db      |||
Qy 540 TGGCGGAGCTCAATTACATTCCTCCCAACCGCGTGGCAACAACCTGGCGGCAACAGTCGT 481
Db      |||
Qy 265 TGCTGATTTGGCGTTGCCACCTCCAGTCTGCGCCCTTCACGCGCCGTCGCAAAATGTCGCGG 324
Db      |||
Qy 480 TGCTGATTTGGCGTTGCCACCTCCAGTCTGCGCCCTTCACGCGCCGTCGCAAAATGTCGCGG 421
Db      |||
Qy 325 CGATTAAATCTCGCGCCGATCAACTGGGTGCGAGGTGGTGTGTCGATGTTAGAACGAA 384
Db      |||
Qy 420 CGATTAAATCTCGCGCCGATCAACTGGGTGCGAGGTGGTGTGTCGATGTTAGAACGAA 361
Db      |||
Qy 385 CGCGCTCGAAGCTGTAAAGCGCGGTGCACAAATCTTCTCGCGCAACGCGTCAGTGGGC 444
Db      |||
Qy 360 CGCGCTCGAAGCTGTAAAGCGCGGTGCACAAATCTTCTCGCGCAACGCGTCAGTGGGC 301
Db      |||
Qy 445 TGATCAATTAATATCCGCTGGATGACCAAGATGTCATTGCTGTGGAAGCTGCCTGCACTA 504
Db      |||
Qy 300 TGATCAATTAATATCCGCTGGATGACCAAGATGTCATTGCTGTGGAAGCTGCCTGCACTA 241
Db      |||
Qy 505 ATGTTCCGGCGGTTATTTCTTGTGATGTCTCTGACGACACCCATCAACAGTATTTTCT 564
Db      |||
```

```

Db      240  ATGTTCCGGGTTATTCTTGATGTCCTGACACAGACCCATCAACAGTATTATTTCT 181
QY      565  CCATGAAGACGGTAGCGACTGGCGGTGGAGCATCTGTCGATTTGGGTACACGACAAA 624
Db      180  CCATGAAGACGGTAGCGACTGGCGGTGGAGCATCTGTCGATTTGGGTACACGACAAA 121
QY      625  TCGCGCTGTTAGCGGCGCCCAATTAAGTTCTGTCGCGCGGTCCTGCGTCTGGCTGGCTGGC 684
Db      120  TCGCGCTGTTAGCGGCGCCCAATTAAGTTCTGTCGCGCGGTCCTGCGTCTGGCTGGCTGGC 62
QY      685  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGACTGGAGTG 744
Db      61  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGACTGGAGTG 744
QY      745  C 745
Db      1  C 1

RESULT 10
BJ063872/c
LOCUS
DEFINITION
BJ063872 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL07k02 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 641)
Email: tshini@genes.nig.ac.jp
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
Kohara,Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
The information of this clone is available through the following
URL.
http://xenopus.nibb.ac.jp.
FEATURES
Location/Qualifiers
1..641
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL07k02"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
ORIGIN
Query Match 19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred. No. 1.8e-177;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 30 CGCGGTATGCATGATAGCCCGGAGAGAGATCAATTACGGTGTGTAATGGAACCA 89
Db 641 CGCGGTATGCATGATAGCCCGGAGAGAGATCAATTACGGTGTGTAATGGAACCA 582
QY 90 GTAAAGTTATACGATGTGCGAGAGTATGCCGGGTCTCTTTATCAGACCGTTTCCCGCGTG 149
Db 581 GTAAAGTTATACGATGTGCGAGAGTATGCCGGGTCTCTTTATCAGACCGTTTCCCGCGTG 522
QY 150 GTGAACCGACCCAGCCAGCTTTCTCGGAACACCGGGGAAAAAGTGAAGCGCGATGGCG 209

```

```

Db      521  GTGAACAGGCGACGCCAGTTTCTCGAAAAACGCGGAAAAAGTGAAGCGCGATGGCG 462
QY      210  GAGCTGAATTACATTCCCAACCGGTGGCAACAACCTGGCGGGCAACAGTCTGCTG 269
Db      461  GAGCTGAATTACATTCCCAACCGGTGGCAACAACCTGGCGGGCAACAGTCTGCTG 402
QY      270  ATTGGCGTTGCACTCAAGTCTGGCCCTGCAACGCGCGCTCGCAAAATTGTCGCGCGGAT 329
Db      401  ATTGGCGTTGCACTCAAGTCTGGCCCTGCAACGCGCGCTCGCAAAATTGTCGCGCGGAT 342
QY      330  AAATCTCGCGCGGATCAACTGGGTGCCAGCGTGGTGTGTCGATGCTAGAACGAGCGGC 389
Db      341  AAATCTCGCGCGGATCAACTGGGTGCCAGCGTGGTGTGTCGATGCTAGAACGAGCGGC 282
QY      390  GTCGAAGCCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 449
Db      281  GTCGAAGCCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 222
QY      450  ATTAACCTATCGCTGGATGACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTAATGTT 509
Db      221  ATTAACCTATCGCTGGATGACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTAATGTT 162
QY      510  CCGCGGTTATTTCTTGATGTCCTGACCAACACCCATCAACAGTATTATTTCTCCCAT 569
Db      161  CCGCGGTTATTTCTTGATGTCCTGACCAACACCCATCAACAGTATTATTTCTCCCAT 102
QY      570  GAAGACGGTACCGGACTGGGCGTGGAGCATCTGTCGCAATGGGTGCCAGCAAAATCGCG 629
Db      101  GAAGACGGTACCGGACTGGGCGTGGAGCATCTGTCGCAATGGGTGCCAGCAAAATCGCG 42
QY      630  CTGTTAGCGGCGCCATTAAGTTCTGTCGCGCGCTCGG 670
Db      41  CTGTTAGCGGCGCCATTAAGTTCTGTCGCGCGCTCGG 1

RESULT 11
CB863814/c
LOCUS
DEFINITION
HH07B20y HH Hordeum vulgare cDNA clone HH07B20 3-PRIME, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 624)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
Graner,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 7 row: B column: 20
Seq primer: SP6.
Location/Qualifiers
1..624
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="Sloop"
/db_xref="GABI:555471"
/db_xref="taxon:4513"
/clone="HH07B20"
/tissue_type="coleoptile"

```

/dev_stage="coleoptile, 1 day old"
/lab_host="DH10B"
/clone_lib="HH"
/note="Vector: pSPORT; Site 1: SalI (5-end of cDNA);
Site 2: NotI (3-end of cDNA); Due to the cloning system
used blue/white selection for recombinants is not 100 %
reliable. Average insert size is 1.3 kb."

ORIGIN

Query Match 18.7%; Score 615.8; DB 6; Length 624;
Best Local Similarity 99.7%; Pred. No. 5.1e-170;
Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GACACATCGAATGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGT 62
DB 619 GACACATCGAATGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGT 560

QY 63 CAATTCAGGGTGGTCAATGTGAACACAGTAAAGTTATACGATGTCGACAGATATCCCGGT 122
DB 559 CAATTCAGGGTGGTCAATGTGAACACAGTAAAGTTATACGATGTCGACAGATATCCCGGT 500

QY 123 GTCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGCGCCAGCCAGTTCCTCGAAAAACG 182
DB 499 GTCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGCGCCAGCCAGTTCCTCGAAAAACG 440

QY 183 CGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCATCCCAACCGGTGSCACAA 242
DB 439 CGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCATCCCAACCGGTGSCACAA 380

QY 243 CAACCTGGCGGCAACACAGTCGTTGCTGATGGCGTTGCCACCTCCAGTCTGCCCTGCAC 302
DB 379 CAACCTGGCGGCAACACAGTCGTTGCTGATGGCGTTGCCACCTCCAGTCTGCCCTGCAC 320

QY 303 GCGCGTGCCTAAATGTGCGCGCATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 362
DB 319 GCGCGTGCCTAAATGTGCGCGCATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 260

QY 363 GTGGTGTGATGGTGAAGACGAAGCGCGTGAAGCTGTAAAGCGCGGTGCACAACTTT 422
DB 259 GTGGTGTGATGGTGAAGACGAAGCGCGTGAAGCTGTAAAGCGCGGTGCACAACTTT 200

QY 423 CTCGGCGAACGGTCAGTGGGTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAGTCCCAT 482
DB 199 CTCGGCGAACGGTCAGTGGGTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAGTCCCAT 140

QY 483 GCTGTGGAAGCTGCTGCACCTAATGTTCCGGGTTATTTCTTGTGCTCTCACACAGACA 542
DB 139 GCTGTGGAAGCTGCTGCACCTAATGTTCCGGGTTATTTCTTGTGCTCTCACACAGACA 80

QY 543 CCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATCTG 602
DB 79 CCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATCTG 20

QY 603 GTCGCAATGGGTCAACAGC 621
DB 19 GTCGCAATGGGTCAACAGC 1

RESULT 12
AL045353/c
LOCUS
DEFINITION DKF2p434B075 s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045353
VERSION AL045353.1 GI:5433506
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 794)
AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann,S.

TITLE EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); sequenced by Qiagen within the cDNA
sequencing consortium of the German Genome Project. r1 sequence
also available.
This clone is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source

1. 794
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p434B075"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 18.5%; Score 612; DB 1; Length 794;
Best Local Similarity 99.7%; Pred. No. 7.3e-169;
Matches 634; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 569 TGAAGACGGTACGGACATGGGCGTGGAGCATCTGTCGCATTTGGTCAACGCAATCGC 628
DB 794 TGAAGACGGTACGGACATGGGCGTGGAG-ATCTGTCGCATTTGGTCAACGCAATCGC 736

QY 629 GCTGTTAGCGGCGCATTAAGTTCCTCGCGCGTCTGCTGCTGGTGGTGGTGGCATAA 688
DB 735 GCTGTTAGCGGCGCATTAAGTTCCTCGCGCGTCTG-CTCGGCTGGTGGCATAA 677

QY 689 ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGCATGAGATGCCAT 748
DB 676 ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGCATGAGATGCCAT 617

QY 749 GTCGCGTTTCAACAAACCATGCAAAATGCTGAATGAGGGCATCGTTCCCATCGCATGCT 808
DB 616 GTCGCGTTTCAACAAACCATGCAAAATGCTGAATGAGGGCATCGTTCCCATCGCATGCT 557

QY 809 GGTTCGCCAACCATCAGATGGCGCTGGGCGCAATCGCGCCATTCACCGATCCGGCTGCG 868
DB 556 GGTTCGCCAACCATCAGATGGCGCTGGGCGCAATTCACCGATCCGGCTGCG 497

QY 869 GTTGGTCCGATATCTCGGTAGTGGGATAGCATACCGAAGACAGCTCATGTTATAT 928
DB 496 GTTGGTCCGATATCTCGGTAGTGGGATAGCATACCGAAGACAGCTCATGTTATAT 437

QY 929 CCCGCGTTAACCAACCATCAAAACAGGATTTTCGCTCTGGGGCAAAACAGCGTGGACCG 988
DB 436 CCCGCGTTAACCAACCATCAAAACAGGATTTTCGCTCTGGGGCAAAACAGCGTGGACCG 377

QY 989 CTTCGTCAACTCTCTCAGGCGCAGCGGTGAGGGCAATCAGCTGTTGCCCGCTCTCAT 1048
DB 376 CTTCGTCAACTCTCTCAGGCGCAGCGGTGAGGGCAATCAGCTGTTGCCCGCTCTCAT 317

QY 1049 GGTGAAAAAGAAAAACCACTGGCGCCCAATACAGCAAAACCGCTCTCCCGCGCGTTGGC 1108
DB 316 GGTGAAAAAGAAAAACCACTGGCGCCCAATACAGCAAAACCGCTCTCCCGCGCGTTGGC 257

QY 1109 CGATTCAATTAATGACGCTGGCACGAGTTTCCCGACTGGAAACGGGCGAGTAGCGCA 1168
DB 256 CGATTCAATTAATGACGCTGGCACGAGTTTCCCGACTGGAAACGGGCGAGTAGCGCA 197

QY 1169 ACGCAATTAATGTGAGTTAGTCTCACTATTAGGCAC 1204
DB 196 ACGCAATTAATGTGAGTTAGTCTCACTATTAGGCAC 161

```

RESULT 13
CL655575/c
LOCUS
DEFINITION
  CL655575
  PRI0124a.C01 - PRI0124a.B21 (752) Mixed stage fosmid library of P.
  pacificus var. California Pristionchus pacificus genomic, genomic
  survey sequence.
  CL655575
  CL655575.1 GI:50134897
  GSS.
  Pristionchus pacificus
  Pristionchus pacificus
  Pristionchus pacificus
  Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
  Neodiplogasteridae; Pristionchus.
  1 (bases 1 to 752)
  Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  AppADB: an AcedB database for the nematode satellite organism
  Pristionchus pacificus
  Nucleic Acids Res. 32 (1), D421-D422 (2004)
  Contact: Sommer RJ
  Evolutionary Biology
  Max-Planck-Institute for Developmental Biology
  Spemannstr. 37-39, Tuebingen D-72076, Germany
  Tel: 00497071601371
  Fax: 00497071601498
  Email: raif.sommer@tuebingen.mpg.de
  This library was generated at Caltech, Pasadena, USA and end
  sequenced at Vancouver, Canada.
  Seq primer: 17
  Class: fosmid ends.
  Location/Qualifiers
    1..752
    /organism="Pristionchus pacificus"
    /mol_type="genomic DNA"
    /strain="California"
    /db_xref="taxon:54126"
    /clone_lib="Mixed stage fosmid library of P. pacificus
    var. California"
    /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
  Query Match      18.3%; Score 603; DB 9; Length 752;
  Best Local Similarity 98.9%; Pred. No. 3.3e-166;
  Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 1996 TTCACCTGGCGCTGATTGCTGCTGAC-GGGGGTTATGCGTTCAAGTATGAAACGGCAA 2054
Db 752 TTCACCTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAACGGCAA 693
QY 2055 GTACGACATTAAAGACGTGGCGGTGGATAACGCTGGCGCGAAACGGGTCTGACCTTCCT 2114
Db 692 GTACGACATTAAAGACGTGGCGGTGGATAACGCTGGCGCGAAACGGGTCTGACCTTCCT 633
QY 2115 GGTTCACCTGATTAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAC 2174
Db 632 GGTTCACCTGATTAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAC 573
QY 2175 TGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGTGGCGATGGTCCACAT 2234
Db 572 TGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGTGGCGATGGTCCACAT 513
QY 2235 CGACACAGCAAGTGAATTATGTTGTAACGGTACTGCGACCTTCAAGGGTCAACCATC 2294
Db 512 CGACACAGCAAGTGAATTATGTTGTAACGGTACTGCGACCTTCAAGGGTCAACCATC 453
QY 2295 CAAACCGTTGTTGGCGTGTGACGCGAGGTATTAAACCGCGCAGTCCGAAACAGAGCT 2354
Db 452 CAAACCGTTGTTGGCGTGTGACGCGAGGTATTAAACCGCGCAGTCCGAAACAGAGCT 393
QY 2355 GGCAAAAGAGTTCTTCGAAAACCTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAA-TA 2413
Db 392 GGCAAAAGAGTTCTTCGAAAACCTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATTA 333

```

```

QY 2414 AAGCAAAACCGCTGGTGCCCTAGCGCTGAAGTCTTACGAGAAAGAGTTGCGGAAAGATC 2473
Db 332 AGGCCAAACCCGTTGGTGCCGCTAGCGCTGAAGTCTTACGAGAAAGAGTTGCGGAAAGATC 273
QY 2474 CACGTATTGCCCGCACCATGAAACCCCAAGAGTGAATCATGCCGAACATCCCGC 2533
Db 272 CACGTATTGCCCGCACCATGAAACCCCAAGAGTGAATCATGCCGAACATCCCGC 213
QY 2534 AGATGTCCGCTTTCTGGTATGCCGCTGCTACTGCGGTGATCAACGCCCGCAGCGGTGTC 2593
Db 212 AGATGTCCGCTTTCTGGTATGCCGCTGCTACTGCGGTGATCAACGCCCGCAGCGGTGTC 153
QY 2594 AGATGTGCGATGAAGCCCTGAAGACGCGCAGACT 2628
Db 152 AGACTGTGCGATGAAGCCCTGAAGACGCGCAGACT 118

RESULT 14
AY080094/c
LOCUS
DEFINITION
  AY080094
  Scripps Pier (La Jolla, CA) uncultured virus community
  uncultured marine virus genomic clone SIO51p3D4L, genomic survey
  sequence.
  AY080094
  AY080094.1 GI:24745276
  GSS.
  uncultured marine virus
  uncultured marine virus
  Viruses; environmental samples.
  1 (bases 1 to 608)
  Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
  Mead,D., Azam,F. and Rohwer,F.
  Genomic analysis of uncultured marine viral communities
  Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
  22294988
  MEDLINE
  PUBMED
  12384570
  COMMENT
  Contact: Rohwer F
  Biology Dept.
  San Diego State University
  5500 Campanile Dr, San Diego, CA 92102, USA
  Tel: 6195941336
  Fax: 619595676
  Email: forest@unstroke.sdsu.edu
  Class: shotgun.
  Location/Qualifiers
    1..608
    /organism="uncultured marine virus"
    /mol_type="genomic DNA"
    /db_xref="taxon:186617"
    /clone="SIO51p3D4L"
    /clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
    community"
    /note="Marine viruses were isolated from 200 liters of
    surface seawater using a combination of differential
    filtration and density-dependent gradient centrifugation.
    Linker-amplified shotgun libraries were constructed by
    randomly shearing the total marine viral community DNA,
    end-repairing, ligating dsDNA linkers to the ends, and
    amplifying the fragments using Vent DNA polymerase. The
    resulting fragments were ligated into the pSMART vector
    and electroporated into MC12 cells (Lucigen; Middleton,
    WI)"

ORIGIN
  Query Match      17.9%; Score 589.8; DB 8; Length 608;
  Best Local Similarity 99.7%; Pred. No. 2.5e-162;
  Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 287 CAGTCTGGCCCTGCACGCGCCCGTCGCAAAATTCGCGGCGATTAAATCTCGCGCGATCA 346
Db 608 CAGTCTGGCCCTGCACGCGCCCATCGCAAAATTCGCGGCGATTAAATCTCGCGCGATCA 549

```

```
QY 347 ACTGGGTGCCAGCGTGGTGTCTCGATGTAAGCAAGCGCGTGAAGCGTGAAGC 406
|
Db 548 ACTGGGTGCCAGCGTGGTGTCTCGATGTAAGCAAGCGCGTGAAGCGTGAAGC 489
|
QY 407 GCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCAATACTATCCGCTGGA 466
|
Db 488 GCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCAATACTATCCGCTGGA 429
|
QY 467 TGACCAGATGCCATGCTGTGGGAAGCTGCCCTGCACATAATGTTCCGGGCTTATTTCTTGA 526
|
Db 428 TGACCAGATGCCATGCTGTGGGAAGCTGCCCTGCACATAATGTTCCAGGCTTATTTCTTGA 369
|
QY 527 TGTCTCTGACCAGACACCCATCAACAGATATATTTCTCCCATGAAGACGGTACGCGACT 586
|
Db 368 TGTCTCTGACCAGACACCCATCAACAGATATATTTCTCCCATGAAGACGGTACGCGACT 309
|
QY 587 GGGCGTGGAGCATCTGGTCGATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAT 646
|
Db 308 GGGCGTGGAGCATCTGGTCGATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAT 249
|
QY 647 AAGTTCTCTCGGCGCTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGT 706
|
Db 248 AAGTTCTCTCGGCGCTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGT 189
|
QY 707 AATTACGCGGATAGCGGAACGGAAGCGACTGGAGTGCCATGTCGGGTTTTTCAACAAC 766
|
Db 188 AATTACGCGGATAGCGGAACGGAAGCGACTGGAGTGCCATGTCGGGTTTTTCAACAAC 129
|
QY 767 CATGCAAAATGCTGAATGAGGCGCATGTTCCCACTGCGATGCTGTTGCAACGATCAGAT 826
|
Db 128 CATGCAAAATGCTGAATGAGGCGCATGTTCCCACTGCGATGCTGTTGCAACGATCAGAT 69
|
QY 827 GCGCTGGGCGCAATGCGGCGCATTTACCGAGTCCGGGCTGCGCGTGTGTCGG 879
|
Db 68 GCGCTGGGCGCAATGCGGCGCATTTACCGAGTCCGGGCTGCGCGTGTGTCGG 16
|
```

```
RESULT 15
AL037742/c
LOCUS
DEFINITION
KFZp564C157.r1.564 (synonym: hfbr2) Homo sapiens cDNA clone
KFZp564C157.5', mRNA sequence.
AL037742
AL037742.1 GI:5407088
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 589)
Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Bloecker, et al.)
Unpublished (1999)
Contact: MIPS
```

```
REFERENCES
AUTHORS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No s1 sequence available.
This clone (DKFZp564C157) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp564C157"
```

```
FEATURES
source
```

```
/tissue_type="brain"
/dev_stages="fetal"
/lab_host="xl-2blue"
/clone_lib="564 (synonym: hfbr2)"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 17.8%; Score 586.4; DB 1; Length 589;
Best Local Similarity 99.7%; Pred. No. 2.5e-161;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 517 TATTTCCTTGATGTCCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACG 576
|
Db 589 TATTTCCTTGATGTCCTGACCAGACACCCATCAACAGTATTATTTTCTCCCTTGAAGACG 530
|
QY 577 GTACGCGATCGGGCGTGGAGCATCTGTGCGATTGGGTCAACAGCAAAATCGCGCTGTTAG 636
|
Db 529 GTACGCGATCGGGCGTGGAGCATCTGTGTCGATTTGGGTCAACAGCAAAATCGCGCTGTTAG 470
|
QY 637 CGGGCCCATTAAGTTCTGTCTCGGGCGCTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGT 696
|
Db 469 CGGGCCCATTAAGTTCTGTCTCGGGCGCTCTGCGTCTGCGTCTGCGTCTGCGTCTGCGT 410
|
QY 697 CTGCAATCAAAATTCAGCCGATAGCGGAACGGAAGCGACTGGAGTGCCATGTCGCGTT 756
|
Db 409 CTGCAATCAAAATTCAGCCGATAGCGGAACGGAAGCGACTGGAGTGCCATGTCGCGTT 350
|
QY 757 TTCAACAAACCATGCAATGCTGAATGAGGCGCATCGTTCCCACTGCGATGCTGTTGCCA 816
|
Db 349 TTCAACAAACCATGCAATGCTGAATGAGGCGCATCGTTCCCACTGCGATGCTGTTGCCA 290
|
QY 817 ACGATCAGATCGGCGTGGGCGCAATCGCGCGCATTTACCGAGTCCGGGCTGCGCGTGGTG 876
|
Db 289 ACGATCAGATCGGCGTGGGCGCAATGCGCGCATTTACCGAGTCCGGGCTGCGCGTGGTG 230
|
QY 877 CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT 936
|
Db 229 CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT 170
|
QY 937 TAACCAACCATCAAAACAGGATTTTCGCTGCTGGGCGCAACACAGCGTGGACCGCTTGTCTGC 996
|
Db 169 TAACCAACCATCAAAACAGGATTTTCGCTGCTGGGCGCAACACAGCGTGGACCGCTTGTCTGC 110
|
QY 997 AACTCTCTCAGGGCGAGCGCGTGAAGGGCAATCAGCTGTTGCCCGCTCTCAGCTGGTGA 1056
|
Db 109 AACTCTCTCAGGGCGAGCGCGTGAAGGGCAATCAGCTGTTGCCCGCTCTCAGCTGGTGA 50
|
QY 1057 GAAAAACCAACCTCGCGCCCAATACGCAAAACCGGCTCTCCCGCGCGTT 1105
|
Db 49 GAAAAACCAACCTCGCGCCCAATACGCAAAACCGGCTCTCCCGCGCGTT 1
```

Search completed: August 3, 2005, 02:02:52
Job time : 6745.2 secs

THIS PAGE BLANK (USPTO)